

Tue Nov 16 16:09:49 2004

us-10-088-666-1.olig10.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:28:49 ; Search time 459.432 Seconds
(without alignments)
3138.794 Million cell updates/sec

Title: US-10-088-666-1

Perfect score: 267

Sequence: 1 tatatggaagtaagaccct.....agtagtggggatcgcccc 267

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 3625171 seqs, 2700493622 residues

Word-size: 10

Total number of hits satisfying chosen parameters: 796812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US6_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	16.5	3989	9	US-09-070-927A-359
2	44	16.5	9797	9	US-09-070-927A-350
3	44	16.5	22960	9	US-09-070-927A-345
4	35	13.1	54	8	US-08-781-986A-5036
5	35	13.1	54	16	US-10-329-624-5036
6	35	13.1	187	8	US-08-781-986A-4682
7	35	13.1	187	16	US-10-329-624-4682
8	35	13.1	216	8	US-08-781-986A-4589
9	35	13.1	216	16	US-10-329-624-4589
10	35	13.1	221	8	US-08-781-986A-4572
11	35	13.1	221	16	US-10-329-624-4572
12	35	13.1	235	8	US-08-781-986A-4536

13	35	13.1	235	16	US-10-329-624-4536	Sequence 4536, Ap
14	35	13.1	242	16	US-08-781-986A-4538	Sequence 4538, Ap
15	35	13.1	242	16	US-10-329-624-4538	Sequence 4538, Ap
16	35	13.1	283	8	US-08-781-986A-4460	Sequence 4460, Ap
17	35	13.1	283	16	US-10-329-624-4460	Sequence 4460, Ap
18	35	13.1	295	8	US-08-781-986A-4440	Sequence 4440, Ap
19	35	13.1	295	16	US-10-329-624-4440	Sequence 4440, Ap
20	35	13.1	300	8	US-08-781-986A-4415	Sequence 4415, Ap
21	35	13.1	300	16	US-10-329-624-4415	Sequence 4415, Ap
22	35	13.1	309	8	US-08-781-986A-4336	Sequence 4336, Ap
23	35	13.1	309	16	US-10-329-624-4336	Sequence 4336, Ap
24	35	13.1	327	8	US-08-781-986A-4342	Sequence 4342, Ap
25	35	13.1	327	16	US-10-329-624-4342	Sequence 4342, Ap
26	35	13.1	330	8	US-08-781-986A-4324	Sequence 4324, Ap
27	35	13.1	330	16	US-10-329-624-4324	Sequence 4324, Ap
28	35	13.1	338	8	US-08-781-986A-4246	Sequence 4246, Ap
29	35	13.1	338	16	US-10-329-624-4246	Sequence 4246, Ap
30	35	13.1	340	8	US-08-781-986A-4195	Sequence 4195, Ap
31	35	13.1	340	16	US-10-329-624-4195	Sequence 4195, Ap
32	35	13.1	348	8	US-08-781-986A-4175	Sequence 4175, Ap
33	35	13.1	348	16	US-10-329-624-4175	Sequence 4175, Ap
34	35	13.1	361	8	US-08-781-986A-3947	Sequence 3947, Ap
35	35	13.1	361	16	US-10-329-624-3947	Sequence 3947, Ap
36	35	13.1	361	16	US-10-329-624-3947	Sequence 3947, Ap
37	35	13.1	367	8	US-08-781-986A-4059	Sequence 4059, Ap
38	35	13.1	367	16	US-10-329-624-4059	Sequence 4059, Ap
39	35	13.1	367	8	US-08-781-986A-4219	Sequence 4219, Ap
40	35	13.1	367	16	US-10-329-624-4059	Sequence 4059, Ap
41	35	13.1	367	16	US-10-329-624-4219	Sequence 4219, Ap
42	35	13.1	370	8	US-08-781-986A-4270	Sequence 4270, Ap
43	35	13.1	370	16	US-10-329-624-4270	Sequence 4270, Ap
44	35	13.1	371	8	US-08-781-986A-4118	Sequence 4118, Ap
45	35	13.1	371	16	US-10-329-624-4118	Sequence 4118, Ap
46	35	13.1	386	8	US-08-781-986A-4064	Sequence 4064, Ap
47	35	13.1	386	16	US-10-329-624-4064	Sequence 4064, Ap
48	35	13.1	389	8	US-08-781-986A-3890	Sequence 3890, Ap
49	35	13.1	389	16	US-10-329-624-3890	Sequence 3890, Ap
50	35	13.1	400	8	US-08-781-986A-3611	Sequence 3611, Ap
51	35	13.1	400	16	US-10-329-624-3611	Sequence 3611, Ap
52	35	13.1	400	8	US-08-781-986A-3634	Sequence 3634, Ap
53	35	13.1	400	16	US-10-329-624-3634	Sequence 3634, Ap
54	35	13.1	400	8	US-08-781-986A-3638	Sequence 3638, Ap
55	35	13.1	400	16	US-10-329-624-3638	Sequence 3638, Ap
56	35	13.1	400	8	US-08-781-986A-3645	Sequence 3645, Ap
57	35	13.1	400	16	US-10-329-624-3645	Sequence 3645, Ap
58	35	13.1	400	8	US-08-781-986A-3670	Sequence 3670, Ap
59	35	13.1	400	16	US-10-329-624-3670	Sequence 3670, Ap
60	35	13.1	400	8	US-08-781-986A-3719	Sequence 3719, Ap
61	35	13.1	400	16	US-10-329-624-3719	Sequence 3719, Ap
62	35	13.1	400	8	US-08-781-986A-3742	Sequence 3742, Ap
63	35	13.1	400	16	US-10-329-624-3742	Sequence 3742, Ap
64	35	13.1	400	8	US-08-781-986A-3748	Sequence 3748, Ap
65	35	13.1	400	16	US-10-329-624-3748	Sequence 3748, Ap
66	35	13.1	400	8	US-08-781-986A-3768	Sequence 3768, Ap
67	35	13.1	400	16	US-10-329-624-3768	Sequence 3768, Ap
68	35	13.1	400	8	US-08-781-986A-3856	Sequence 3856, Ap
69	35	13.1	400	16	US-10-329-624-3856	Sequence 3856, Ap
70	35	13.1	400	8	US-08-781-986A-4029	Sequence 4029, Ap
71	35	13.1	400	16	US-10-329-624-4029	Sequence 4029, Ap
72	35	13.1	400	16	US-10-329-624-3634	Sequence 3634, Ap
73	35	13.1	400	16	US-10-329-624-3634	Sequence 3634, Ap
74	35	13.1	400	16	US-10-329-624-3645	Sequence 3645, Ap
75	35	13.1	400	16	US-10-329-624-3645	Sequence 3645, Ap
76	35	13.1	400	16	US-10-329-624-3660	Sequence 3660, Ap
77	35	13.1	400	16	US-10-329-624-3660	Sequence 3660, Ap
78	35	13.1	400	16	US-10-329-624-3671	Sequence 3671, Ap
79	35	13.1	400	16	US-10-329-624-3671	Sequence 3671, Ap
80	35	13.1	400	16	US-10-329-624-3719	Sequence 3719, Ap
81	35	13.1	400	16	US-10-329-624-3719	Sequence 3719, Ap
82	35	13.1	400	16	US-10-329-624-3742	Sequence 3742, Ap
83	35	13.1	400	16	US-10-329-624-3742	Sequence 3742, Ap
84	35	13.1	400	16	US-10-329-624-3748	Sequence 3748, Ap
85	35	13.1	400	16	US-10-329-624-3768	Sequence 3768, Ap

86 35 13.1 400 16 US-10-329-624-3858 Sequence 3858, Ap
C 87 35 13.1 400 16 US-10-329-624-3866 Sequence 3866, Ap
C 88 35 13.1 400 16 US-10-329-624-3934 Sequence 3934, Ap
C 89 35 13.1 400 16 US-10-329-624-4029 Sequence 4029, Ap
C 90 35 13.1 401 6 US-08-781-986A-3731 Sequence 3731, Ap
C 91 35 13.1 401 6 US-10-329-624-3731 Sequence 3731, Ap
C 92 35 13.1 425 8 US-08-781-986A-3714 Sequence 3714, Ap
C 93 35 13.1 425 16 US-10-329-624-3714 Sequence 3714, Ap
C 94 35 13.1 458 8 US-08-781-986A-3757 Sequence 3757, Ap
C 95 35 13.1 458 16 US-10-329-624-3757 Sequence 3757, Ap
C 96 35 13.1 475 8 US-08-781-986A-3621 Sequence 3621, Ap
C 97 35 13.1 475 16 US-10-329-624-3621 Sequence 3621, Ap
C 98 35 13.1 587 8 US-08-781-986A-3554 Sequence 3554, Ap
C 99 35 13.1 587 16 US-10-329-624-3554 Sequence 3554, Ap
C 100 35 13.1 596 8 US-08-781-986A-3596 Sequence 3596, Ap

ALIGNMENTS

RESULT 1
US-09-070-927A-359/c
; Sequence 359, Application US/09070927A
; Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 359:
SEQUENCE CHARACTERISTICS:
LENGTH: 3989 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 359:
US-09-070-927A-359
Query Match 16.5%; Score 44; DB 9; Length 3989;
Best Local Similarity 100.0%; Pred. No. 4,7e-14;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 GAAGATACACCTGTCCTCCATGCCGAACACAGAGTTAAGCTTC 236
Db 3627 GAAGATACACCTGTCCTCCATGCCGAACACAGAGTTAAGCTTC 3584

RESULT 2

US-09-070-927A-550/c
; Sequence 550, Application US/09070927A
; Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 550:
SEQUENCE CHARACTERISTICS:
LENGTH: 9797 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 550:
US-09-070-927A-550
Query Match 16.5%; Score 44; DB 9; Length 9797;
Best Local Similarity 100.0%; Pred. No. 4,7e-14;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 193 GAAGATACACCTGTCCTCCATGCCGAACACAGAGTTAAGCTTC 236
Db 9424 GAAGATACACCTGTCCTCCATGCCGAACACAGAGTTAAGCTTC 9381
RESULT 3
US-09-070-927A-345
; Sequence 345, Application US/09070927A
; Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:18:54 | Search time 113.856 Seconds
(without alignments)
1666.850 Million cell updates/sec

Title: US-10-088-666-1
Perfect score: 267
Sequence: 1 tatatcgaagtaagaccctc.....agtagctg99gagtcgcccc 267

Scoring table: OLIGO_NUC:
Gapop 60.0, Gapext 60.0

Searched: 824507 seqs, 35394441 residues

Word size: 10

Total number of hits satisfying chosen parameters: 100240

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database:

Issued Parents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCRTS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfillseq1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	16.5	249	4 US-09-134-000C-3148	Sequence 3148, Ap
2	35	13.1	54	4 US-08-956-171E-5036	Sequence 5036, Ap
3	35	13.1	54	4 US-08-781-986A-5036	Sequence 5036, Ap
4	35	13.1	187	4 US-08-956-171E-4682	Sequence 4682, Ap
5	35	13.1	187	4 US-08-781-986A-4682	Sequence 4682, Ap
6	35	13.1	216	4 US-08-956-171E-4589	Sequence 4589, Ap
7	35	13.1	216	4 US-08-781-986A-4589	Sequence 4589, Ap
8	35	13.1	221	4 US-08-956-171E-4572	Sequence 4572, Ap
9	35	13.1	221	4 US-08-781-986A-4572	Sequence 4572, Ap
10	35	13.1	235	4 US-08-956-171E-4536	Sequence 4536, Ap
11	35	13.1	235	4 US-08-781-986A-4536	Sequence 4536, Ap
12	35	13.1	242	4 US-08-956-171E-4538	Sequence 4538, Ap
13	35	13.1	242	4 US-08-781-986A-4538	Sequence 4538, Ap
14	35	13.1	283	4 US-08-956-171E-4460	Sequence 4460, Ap
15	35	13.1	283	4 US-08-781-986A-4460	Sequence 4460, Ap
16	35	13.1	295	4 US-08-956-171E-4440	Sequence 4440, Ap
17	35	13.1	295	4 US-08-781-986A-4440	Sequence 4440, Ap
18	35	13.1	300	4 US-08-956-171E-4415	Sequence 4415, Ap
19	35	13.1	300	4 US-08-781-986A-4415	Sequence 4415, Ap
20	35	13.1	309	4 US-08-956-171E-4326	Sequence 4326, Ap
21	35	13.1	309	4 US-08-781-986A-4326	Sequence 4326, Ap
22	35	13.1	327	4 US-08-956-171E-4342	Sequence 4342, Ap
23	35	13.1	327	4 US-08-781-986A-4342	Sequence 4342, Ap
24	35	13.1	330	4 US-08-956-171E-4374	Sequence 4374, Ap
25	35	13.1	330	4 US-08-781-986A-4374	Sequence 4374, Ap
26	35	13.1	338	4 US-08-956-171E-4246	Sequence 4246, Ap
27	35	13.1	338	4 US-08-781-986A-4246	Sequence 4246, Ap

28	35	13.1	340	4 US-08-956-171E-4195	Sequence 4195, Ap
29	35	13.1	340	4 US-08-781-986A-4195	Sequence 4195, Ap
30	35	13.1	348	4 US-08-956-171E-4175	Sequence 4175, Ap
31	35	13.1	348	4 US-08-781-986A-4175	Sequence 4175, Ap
32	35	13.1	361	4 US-08-956-171E-3947	Sequence 3947, Ap
33	35	13.1	361	4 US-08-956-171E-4296	Sequence 4296, Ap
34	35	13.1	361	4 US-08-781-986A-3947	Sequence 3947, Ap
35	35	13.1	361	4 US-08-781-986A-4296	Sequence 4296, Ap
36	35	13.1	367	4 US-08-956-171E-4059	Sequence 4059, Ap
37	35	13.1	367	4 US-08-781-986A-4059	Sequence 4059, Ap
38	35	13.1	367	4 US-08-956-171E-4219	Sequence 4219, Ap
39	35	13.1	367	4 US-08-781-986A-4219	Sequence 4219, Ap
40	35	13.1	370	4 US-08-956-171E-4270	Sequence 4270, Ap
41	35	13.1	370	4 US-08-781-986A-4270	Sequence 4270, Ap
42	35	13.1	371	4 US-08-956-171E-4118	Sequence 4118, Ap
43	35	13.1	371	4 US-08-781-986A-4118	Sequence 4118, Ap
44	35	13.1	386	4 US-08-956-171E-4054	Sequence 4054, Ap
45	35	13.1	386	4 US-08-781-986A-4054	Sequence 4054, Ap
46	35	13.1	389	4 US-08-956-171E-3850	Sequence 3850, Ap
47	35	13.1	389	4 US-08-781-986A-3850	Sequence 3850, Ap
48	35	13.1	389	4 US-08-956-171E-3611	Sequence 3611, Ap
49	35	13.1	400	4 US-08-956-171E-3624	Sequence 3624, Ap
50	35	13.1	400	4 US-08-956-171E-3631	Sequence 3631, Ap
51	35	13.1	400	4 US-08-956-171E-3634	Sequence 3634, Ap
52	35	13.1	400	4 US-08-956-171E-3638	Sequence 3638, Ap
53	35	13.1	400	4 US-08-956-171E-3645	Sequence 3645, Ap
54	35	13.1	400	4 US-08-956-171E-3660	Sequence 3660, Ap
55	35	13.1	400	4 US-08-956-171E-3661	Sequence 3661, Ap
56	35	13.1	400	4 US-08-956-171E-3675	Sequence 3675, Ap
57	35	13.1	400	4 US-08-956-171E-3679	Sequence 3679, Ap
58	35	13.1	400	4 US-08-956-171E-3719	Sequence 3719, Ap
59	35	13.1	400	4 US-08-956-171E-3738	Sequence 3738, Ap
60	35	13.1	400	4 US-08-956-171E-3742	Sequence 3742, Ap
61	35	13.1	400	4 US-08-956-171E-3748	Sequence 3748, Ap
62	35	13.1	400	4 US-08-956-171E-3748	Sequence 3748, Ap
63	35	13.1	400	4 US-08-956-171E-3768	Sequence 3768, Ap
64	35	13.1	400	4 US-08-956-171E-3858	Sequence 3858, Ap
65	35	13.1	400	4 US-08-956-171E-3866	Sequence 3866, Ap
66	35	13.1	400	4 US-08-956-171E-3924	Sequence 3924, Ap
67	35	13.1	400	4 US-08-956-171E-4029	Sequence 4029, Ap
68	35	13.1	400	4 US-08-781-986A-3611	Sequence 3611, Ap
69	35	13.1	400	4 US-08-781-986A-3624	Sequence 3624, Ap
70	35	13.1	400	4 US-08-781-986A-3631	Sequence 3631, Ap
71	35	13.1	400	4 US-08-781-986A-3634	Sequence 3634, Ap
72	35	13.1	400	4 US-08-781-986A-3638	Sequence 3638, Ap
73	35	13.1	400	4 US-08-781-986A-3645	Sequence 3645, Ap
74	35	13.1	400	4 US-08-781-986A-3660	Sequence 3660, Ap
75	35	13.1	400	4 US-08-781-986A-3661	Sequence 3661, Ap
76	35	13.1	400	4 US-08-781-986A-3670	Sequence 3670, Ap
77	35	13.1	400	4 US-08-781-986A-3675	Sequence 3675, Ap
78	35	13.1	400	4 US-08-781-986A-3719	Sequence 3719, Ap
79	35	13.1	400	4 US-08-781-986A-3738	Sequence 3738, Ap
80	35	13.1	400	4 US-08-781-986A-3742	Sequence 3742, Ap
81	35	13.1	400	4 US-08-781-986A-3744	Sequence 3744, Ap
82	35	13.1	400	4 US-08-781-986A-3768	Sequence 3768, Ap
83	35	13.1	400	4 US-08-781-986A-3768	Sequence 3768, Ap
84	35	13.1	400	4 US-08-781-986A-3858	Sequence 3858, Ap
85	35	13.1	400	4 US-08-781-986A-3866	Sequence 3866, Ap
86	35	13.1	400	4 US-08-781-986A-3924	Sequence 3924, Ap
87	35	13.1	400	4 US-08-781-986A-4029	Sequence 4029, Ap
88	35	13.1	401	4 US-08-956-171E-3731	Sequence 3731, Ap
89	35	13.1	401	4 US-08-781-986A-3731	Sequence 3731, Ap
90	35	13.1	425	4 US-08-956-171E-3714	Sequence 3714, Ap
91	35	13.1	425	4 US-08-781-986A-3714	Sequence 3714, Ap
92	35	13.1	458	4 US-08-956-171E-3757	Sequence 3757, Ap
93	35	13.1	458	4 US-08-781-986A-3757	Sequence 3757, Ap
94	35	13.1	475	4 US-08-956-171E-3621	Sequence 3621, Ap
95	35	13.1	475	4 US-08-781-986A-3621	Sequence 3621, Ap
96	35	13.1	587	4 US-08-956-171E-3554	Sequence 3554, Ap
97	35	13.1	587	4 US-08-781-986A-3554	Sequence 3554, Ap
98	35	13.1	596	4 US-08-956-171E-3596	Sequence 3596, Ap
99	35	13.1	596	4 US-08-781-986A-3596	Sequence 3596, Ap
100	35	13.1	840	4 US-08-956-171E-508	Sequence 508, Ap

ALIGNMENTS

RESULT 1

US-09-134-000C-3148
Sequence 3148, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patent version 3.1
SEQ ID NO 3148
LENGTH: 249
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-3148

Query Match 16.5%; Score 44; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 GAAGATACACCTGTTCCATGCCGACACAGAGTTAAGCTTC 236
DB 48 GAAGATACACCTGTTCCATGCCGACACAGAGTTAAGCTTC 91

RESULT 2

US-08-956-171E-5036/C
Sequence 5036, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gill H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Farnon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5236
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5036:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 5036:
US-08-956-171E-5036

Query Match 13.1%; Score 35; DB 4; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 ACACCTGTTCCATGCCGACACAGAGTTAAGCT 234
DB 52 ACACCTGTTCCATGCCGACACAGAGTTAAGCT 18

RESULT 3

US-08-781-986A-5036/C
Sequence 5036, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gill H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Farnon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5036:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-5036

Query Match 13.1%; Score 35; DB 4; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 ACACCTGTTCCATGCCGACACAGAGTTAAGCT 234
DB 52 ACACCTGTTCCATGCCGACACAGAGTTAAGCT 18

RESULT 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:59:59 ; Search time 517.162 Seconds

(without alignments)
2710.166 Million cell updates/sec

Title: US-10-088-666-1

Sequence: 267
1 tatatggaagtaagaccctc.....agtagtgggggagtcgcc 267Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 413486 seqs, 2624710521 residues

Word size: 10

Total number of hits satisfying chosen parameters: 703887

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : N_Geneseq_23sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	267	100.0	267	5	AAFe1566	Lactobaci
2	73	27.3	326	5	AAFe1567	Lactobaci
3	50	18.7	338	5	AAFe1570	Lactobaci
4	50	18.7	351	5	AAFe1568	Lactobaci
5	50	18.7	414	5	AAFe1569	Lactobaci
6	44	16.5	249	10	ADH8563	Enterococ
7	44	16.5	3989	2	AAK13286	Enterococ
8	44	16.5	3989	6	ABS99091	Enterococ
9	44	16.5	9797	6	AAK13487	Enterococ
10	44	16.5	9797	6	ABS99282	Enterococ
11	44	16.5	22960	6	AAK13282	Enterococ
12	44	16.5	22960	6	ABS99077	Enterococ
13	44	16.5	110000	10	ADFe77343_05	Continuation (6 of 6)
14	44	16.5	110000	10	ADFe77343_07	Continuation (8 of 8)
15	44	16.5	110000	10	ADFe77343_18	Continuation (19 of 19)
16	40	15.0	317	5	AAFe1572	L. coryni
17	40	15.0	317	5	AAFe1571	L. coryni
18	40	15.0	326	5	AAFe1575	Pediococc
19	40	15.0	335	5	AAFe1574	Pediococc
20	40	15.0	2942	6	AB154766	Pediococc
21	35	13.1	54	2	AAV79347	Staphyloc

22	35	13.1	187	2	AAV78993 Staphyloc
23	35	13.1	216	2	AAV78900 Staphyloc
24	35	13.1	221	2	AAV78883 Staphyloc
25	35	13.1	235	2	AAV78847 Staphyloc
26	35	13.1	242	2	AAV78849 Staphyloc
27	35	13.1	283	2	AAV78771 Staphyloc
28	35	13.1	300	2	AAV78726 Staphyloc
29	35	13.1	309	2	AAV78637 Staphyloc
30	35	13.1	327	2	AAV78653 Staphyloc
31	35	13.1	330	2	AAV78635 Staphyloc
32	35	13.1	338	2	AAV78557 Staphyloc
33	35	13.1	340	2	AAV78506 Staphyloc
34	35	13.1	348	2	AAV78486 Staphyloc
35	35	13.1	361	2	AAV78258 Staphyloc
36	35	13.1	361	2	AAV78607 Staphyloc
37	35	13.1	367	2	AAV78370 Staphyloc
38	35	13.1	367	2	AAV78530 Staphyloc
39	35	13.1	370	2	AAV78581 Staphyloc
40	35	13.1	371	2	AAV78429 Staphyloc
41	35	13.1	386	2	AAV78375 Staphyloc
42	35	13.1	389	2	AAV78201 Staphyloc
43	35	13.1	400	2	AAV77942 Staphyloc
44	35	13.1	400	2	AAV77981 Staphyloc
45	35	13.1	400	2	AAV77986 Staphyloc
46	35	13.1	400	2	AAV78059 Staphyloc
47	35	13.1	400	2	AAV78340 Staphyloc
48	35	13.1	400	2	AAV77922 Staphyloc
49	35	13.1	400	2	AAV77971 Staphyloc
50	35	13.1	400	2	AAV78049 Staphyloc
51	35	13.1	400	2	AAV78169 Staphyloc
52	35	13.1	400	2	AAV77972 Staphyloc
53	35	13.1	400	2	AAV78055 Staphyloc
54	35	13.1	400	2	AAV78030 Staphyloc
55	35	13.1	400	2	AAV78079 Staphyloc
56	35	13.1	400	2	AAV77945 Staphyloc
57	35	13.1	400	2	AAV78053 Staphyloc
58	35	13.1	400	2	AAV77935 Staphyloc
59	35	13.1	400	2	AAV78245 Staphyloc
60	35	13.1	400	2	AAV77956 Staphyloc
61	35	13.1	400	2	AAV77949 Staphyloc
62	35	13.1	400	2	AAV78177 Staphyloc
63	35	13.1	401	2	AAV78042 Staphyloc
64	35	13.1	425	2	AAV78025 Staphyloc
65	35	13.1	475	2	AAV77932 Staphyloc
66	35	13.1	587	2	AAV77865 Staphyloc
67	35	13.1	596	2	AAV77907 Staphyloc
68	35	13.1	840	2	AAV74819 Staphyloc
69	35	13.1	1051	2	AAV77885 Staphyloc
70	35	13.1	1089	4	AAH54969 S. epider
71	35	13.1	1171	2	AAV77850 Staphyloc
72	35	13.1	2209	2	AAV77863 Staphyloc
73	35	13.1	2662	4	AAH54970 S. epider
74	35	13.1	2725	4	AAH54148 S. epider
75	35	13.1	2869	4	AAH55038 S. epider
76	35	13.1	2912	10	ADBe1699 23S rRNA
77	35	13.1	3012	4	AAH54319 S. epider
78	35	13.1	3012	4	AAH54343 S. epider
79	35	13.1	3019	4	AAH54554 S. epider
80	35	13.1	3030	4	AAH54552 S. epider
81	35	13.1	3232	4	AAH54853 S. epider
82	35	13.1	3305	4	AAH54792 S. epider
83	35	13.1	3332	4	AAH54919 S. epider
84	35	13.1	3405	4	AAH54539 S. epider
85	35	13.1	3608	4	AAH54395 S. epider
86	35	13.1	3625	4	AAH54294 S. epider
87	35	13.1	3656	4	AAH54375 S. epider
88	35	13.1	3732	4	AAH54994 S. epider
89	35	13.1	3845	4	AAH54156 S. epider
90	35	13.1	3937	4	AAH54408 S. epider
91	35	13.1	4106	4	AAH54320 S. epider
92	35	13.1	4346	4	AAH54338 S. epider
93	35	13.1	6421	8	ADH59790 Staphyloc
94	35	13.1	6591	2	AAV77425 Staphyloc

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:16:19 : Search time 1543.47 Seconds
(without alignments)
6303.600 Million cell updates/sec

Title: US-10-088-666-1
Perfect score: 267
Sequence: 1 tatatgaagtaagacccct.....agtagtcggggagtcgcccc 267
Scoring table: Oligo-MUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 segs, 18219865908 residues

Word size: 10

Total number of hits satisfying chosen parameters: 6545258

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	21.0	793	8	BZ669094 C01100.1.
2	22	8.2	3268	8	BH770998 LMGtag72
3	22	8.2	6499	8	BH771024 LMGtag74
4	20	7.5	693	4	BI508715 BH170011A
5	20	7.5	693	4	BZ170405 CH230-369
6	20	7.5	830	8	BZ076210 LK51d10.
7	19	7.1	547	8	BH386373 AG-ND-123
8	19	7.1	586	8	BH823931 BACP2-D0
9	19	7.1	622	8	AQ779183 HS_3185_A
10	19	7.1	667	2	BH865729 BB865729
11	19	7.1	672	2	BH396985 AG-ND-137
12	19	7.1	696	2	BE113195 UI-R-BJ1-
13	19	7.1	703	8	BH367775 AG-ND-137
14	19	7.1	703	8	BH402869 AG-ND-140
15	19	7.1	754	8	BH399988 AG-ND-137
16	19	7.1	759	8	BH381434 AG-ND-137
17	19	7.1	800	5	BX881725 BX881725
18	19	7.1	829	8	AZ540485 ENTG39TF
19	19	7.1	877	8	CC141515 NDL32K2.
20	19	7.1	888	8	CG965204 MBEMD88TF
21	18	6.7	161	1	AA815431 a165a05.s
22	18	6.7	233	5	BQ466564 HT02011r
23	18	6.7	275	4	BX376623 EBem05_SQ
24	18	6.7	286	1	AV155185 AV155185
25	18	6.7	300	1	AJ480648
26	18	6.7	306	2	BR359579
27	18	6.7	309	2	BB116773
28	18	6.7	328	7	CO750077
29	18	6.7	338	1	AU112531
30	18	6.7	356	8	BH402056 AG-ND-119
31	18	6.7	360	8	AJ469080
32	18	6.7	360	1	AJ474156
33	18	6.7	360	1	AJ474753
34	18	6.7	379	8	AQ078388
35	18	6.7	398	2	BE103481
36	18	6.7	398	2	BE211444
37	18	6.7	407	8	AZ167077
38	18	6.7	413	1	AJ436048
39	18	6.7	417	7	CO750063
40	18	6.7	420	1	AJ469076
41	18	6.7	420	1	AJ469077
42	18	6.7	420	1	AJ469078
43	18	6.7	420	1	AJ474752
44	18	6.7	424	2	AM662006
45	18	6.7	426	1	AA815371
46	18	6.7	435	4	BI535214
47	18	6.7	480	1	AJ469079
48	18	6.7	483	4	BM098432
49	18	6.7	499	2	AM318259
50	18	6.7	501	4	BG047555
51	18	6.7	501	5	BQ754077
52	18	6.7	513	6	CB871365
53	18	6.7	548	1	AV930741
54	18	6.7	558	8	AZ216879
55	18	6.7	560	4	BG987436
56	18	6.7	562	2	AM510851
57	18	6.7	562	8	AQ247784 HS_2063.A
58	18	6.7	563	2	BE755776
59	18	6.7	584	5	BQ412586
60	18	6.7	604	2	AM985312
61	18	6.7	607	7	CF668065
62	18	6.7	609	6	CB438313
63	18	6.7	618	7	CO750644
64	18	6.7	620	6	CB863479
65	18	6.7	620	8	AQ476647
66	18	6.7	627	9	BY657134
67	18	6.7	653	8	BH269565
68	18	6.7	668	6	CB006833
69	18	6.7	668	5	CR286805
70	18	6.7	676	6	BY707349
71	18	6.7	684	7	CK974516
72	18	6.7	719	6	CB420219
73	18	6.7	719	6	CB420267
74	18	6.7	723	9	EX178794
75	18	6.7	727	8	AZ409026
76	18	6.7	738	7	CR286927
77	18	6.7	773	4	BI416176
78	18	6.7	773	9	CG072964
79	18	6.7	779	8	BH542735
80	18	6.7	781	8	AZ210199
81	18	6.7	789	9	CG883248
82	18	6.7	797	8	BZ192101
83	18	6.7	805	7	CK73148
84	18	6.7	807	7	CK362822
85	18	6.7	810	7	CR823773
86	18	6.7	848	8	AZ683885
87	18	6.7	872	7	CG118408
88	18	6.7	894	7	CG149988
89	18	6.7	894	7	CG750445
90	18	6.7	898	8	BH133066
91	18	6.7	900	8	BH147215
92	18	6.7	903	8	BH136872
93	18	6.7	937	9	CG118407
94	18	6.7	943	8	AZ136157
95	18	6.7	959	9	CG609999
96	18	6.7	985	9	CG610004
97	18	6.7	1227	9	CG751219
AJ480648	AJ480648				
BR359579	BR359579				
BB116773	BB116773				
CO750077	KBFAE004P				
AU112531	AU112531				
BH402056	AG-ND-119				
AJ469080	AJ469080				
AJ474156	AJ474156				
AJ474753	AJ474753				
AQ078388	CIT-ASP-2				
BE103481	UI-R-BX0-				
BE211444	SO40512.Y				
AZ167077	SP 0142.B				
AJ436048	AJ436048				
CO750063	KBFAE004O				
AJ469076	AJ469076				
AJ469077	AJ469077				
AJ469078	AJ469078				
AJ474752	AJ474752				
AM662006	h179n08.x				
AA815371	a164906.s				
BI535214	398595.MA				
AJ469079	AJ469079				
BM098432	EBem08_SQ				
AM318259	sg62h12.Y				
BG047555	aa72h09.Y				
BQ754077	EBGA01.SQ				
CB871365	HC03101Y				
AV930741	AV930741				
AZ216879	Sheared.D				
BG987436	MRI-HT106				
AM510851	h440e02.x				
AQ247784	HS_2063.A				
BE755776	209704.MA				
BQ412586	GA_EB005				
AM985312	NXNV_135				
CF668065	RTCT1_34				
CB438313	66864.MA				
CO750644	KBFAE007N				
CB863479	HH04n11w				
AQ476647	CITB1-EL-				
BY657134	Reverse.s				
BH269565	CH230-122				
CB006833	VVC038E01				
CR286805	CR286805				
BY707349	BY707349				
CK974516	4105025.B				
CB420219	593151.MA				
CB420267	593204.MA				
EX178794	Dario.fet				
AZ409026	IN0180E22				
CR286927	CR286927				
BI416176	602989066				
CG072964	PIEXE76TD				
BH542735	BOGXU95TR				
AZ210199	SP 0153.A				
CG883248	ZKMBB045				
BZ192101	CH230-279				
CK73148	961754.MA				
CK362822	AGENCOURT				
CR823773	Oa.sp01n				
AZ683885	ENTLP32TF				
CG118408	PUPOH73TB				
CG149988	KBFAE004G				
CG750445	KATFAE007C				
BH133066	ENTPAC6TR				
BH147215	ENTOP41TR				
BH136872	BNTAB2TR				
CG118407	PUPOH73TB				
AZ136157	SP 0169.A				
CG609999	OGTIN26TH				
CG610004	OGULH26TV				
CG751219	P045-3-G1				

98 6.7 1258 8 BH770957 LMGtag68
c 99 6.7 1344 9 AG031219
100 6.4 112 9 CG544484 OST141740

ALIGNMENTS

RESULT 1
LOCUS BZ369094 793 bp DNA linear GSS 19-NOV-2002
DEFINITION Cot100.1.2.D04 Maize Cot100 library Zea mays genomic, genomic survey sequence.

ACCESSION BZ369094
VERSION BZ369094
KEYWORDS GI:25122715

SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Bukariyola, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 793)

AUTHORS Yuan.Y., Samignuel.P. and Bennetzen.J.L.
TITLE High Cot sequence analysis of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Bennetzen JL
Department of Biological Sciences
Purdue University
Hansen 339F, Purdue University, West Lafayette, IN 47907, USA
Tel: 765 494 4419
Fax: 765 496 1496
Email: maize@ilbo.bio.purdue.edu

Forward and reverse reads were assembled when significant overlap was detected.
Seg primer: T7 and T3
Class: shotgun.

FEATURES
source location/Qualifiers

1..793
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/dev_stage="10 days seedling"
/lab_host="PH10B"
/lab_host="Maize Cot100 library"
/note="Organ: young leaves; Vector: pcr4TOPO, Maize genomic DNA was sheared to fragments averaging about 1.8 kb, was denatured and then reassociated in 1 X SSC at 65°C. After a given Cot value was reached, aliquots were run over a hydroxyapatite (HAP) column in order to separate single stranded DNA from double stranded DNA. The single stranded DNA was then converted to a double stranded form with one round of Klenow DNA polymerase treatment with random 9-mer primers. The double-stranded fragments were then further size-selected over an agarose gel and cloned into the PCR4 TOPO vector."

ORIGIN

Query Match 21.0%; Score 56; DB 8; Length 793;
Best Local Similarity 100.0%; Freq. No. 1.4e-20; Indels 0; Gaps 0;
Matches 56; Conservative 0; Mismatches 0;

OY 189 GCTGAAGATACACCTGTTCCATGCCAGACAGAAATTAGCTTACGACGCC 244
DB 314 GCTGAAGATACACCTGTTCCATGCCAGACAGAAATTAGCTTACGACGCC 369

RESULT 2
LOCUS BH770998 3268 bp DNA linear GSS 01-MAY-2002
DEFINITION LMGtag721 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, genomic survey sequence.
ACCESSION BH770998

VERSION BH770998.1 GI:20373955
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris
ORGANISM Lactococcus lactis subsp. cremoris

REFERENCE 1 (bases 1 to 3268)
AUTHORS Bolotin.A., Ehrlich.S.D. and Sorokin.A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Aliments (2002) In press
COMMENT Contact: Sorokin A
Genetique Microbienne

INRA CNJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21

Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is dnag (93%)

Class: shotgun
High quality sequence start: 30
High quality sequence stop: 3240.

FEATURES
source location/Qualifiers

1..3268
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MG1363"
/sub_species="cremoris"
/db_xref="taxon:1359"
/clone_lib="MG1363 Random Sequence Tag Library"
/note="Vector: pSGMW2; Site 1: Small; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AulI digestion or by sonication."

ORIGIN

Query Match 8.2%; Score 22; DB 8; Length 3268;
Best Local Similarity 100.0%; Freq. No. 0.51;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 TGGAGCGGACCACTACTAATCG 89
DB 422 TGGAGCGGACCACTACTAATCG 443

RESULT 3
LOCUS BH771024 6499 bp DNA linear GSS 01-MAY-2002
DEFINITION LMGtag746 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, genomic survey sequence.
ACCESSION BH771024
VERSION BH771024.1 GI:20373981

KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris
ORGANISM Lactococcus lactis subsp. cremoris

REFERENCE 1 (bases 1 to 6499)

AUTHORS Bolotin.A., Ehrlich.S.D. and Sorokin.A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Aliments (2002) In press
COMMENT Contact: Sorokin A
Genetique Microbienne

INRA CNJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21

Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is ywga (78%)

Class: shotgun
High quality sequence start: 30
High quality sequence stop: 6471.

FEATURES
source location/Qualifiers
1..6499
/organism="Lactococcus lactis subsp. cremoris"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:16:19 ; Search time 115.616 Seconds
(without alignments)
6303.600 Million cell updates/sec

Title: US-10-088-666-74

Perfect score: 20

Sequence: 1 gagggagaagatctctat 20

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219855908 residues

Word size: 10

Total number of hits satisfying chosen parameters: 725526

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : EST:

1: gb_est1:*
2: gb_est2:*
3: gb_hnc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	95.0	613	4	BI508715
2	17	85.0	268	4	BA34970
3	17	85.0	296	4	BI296803
4	17	85.0	584	1	AU016700
5	17	85.0	696	1	BE113195
6	17	85.0	1025	9	CL452767
7	16	80.0	205	4	BF998912
8	16	80.0	262	8	ACQ471789
9	16	80.0	301	2	BA36559
10	16	80.0	409	1	AV540606
11	16	80.0	411	6	AZ232723
12	16	80.0	419	6	CB764501
13	16	80.0	425	1	AA317772
14	16	80.0	425	8	AC0860424
15	16	80.0	442	8	CL658098
16	16	80.0	445	6	CB744432
17	16	80.0	454	5	BO832471
18	16	80.0	467	6	CB714203
19	16	80.0	468	1	AL046182
20	16	80.0	520	9	CE570149
21	16	80.0	531	6	CA630367
22	16	80.0	542	4	BJ249708
23	16	80.0	573	9	CE251105
24	16	80.0	575	2	BF483725

25	16	80.0	581	8	AZ982290
26	16	80.0	590	4	BM704665
27	16	80.0	595	6	CB177816
28	16	80.0	607	4	BJ567993
29	16	80.0	607	6	CB634129
30	16	80.0	620	5	BX254803
31	16	80.0	635	8	AZ384762
32	16	80.0	650	4	BI113486
33	16	80.0	658	7	CN305165
34	16	80.0	662	5	BO829603
35	16	80.0	671	6	CB840482
36	16	80.0	671	5	BX918106
37	16	80.0	674	6	CB841003
38	16	80.0	682	9	AG295673
39	16	80.0	706	4	BI221782
40	16	80.0	730	8	AZ342561
41	16	80.0	737	9	AG422832
42	16	80.0	748	4	BG921699
43	16	80.0	762	2	BE572798
44	16	80.0	765	4	BI869712
45	16	80.0	766	7	CO397391
46	16	80.0	812	8	BH096352
47	16	80.0	881	2	BF123606
48	16	80.0	882	9	CL069801
49	16	80.0	958	4	BF344853
50	16	80.0	982	5	BO950615
51	15	75.0	92	6	CD838172
52	15	75.0	141	1	AA795871
53	15	75.0	150	5	BX511574
54	15	75.0	165	7	CF652072
55	15	75.0	172	5	BP015388
56	15	75.0	180	2	BA35135
57	15	75.0	182	2	AW616050
58	15	75.0	182	2	BA35118
59	15	75.0	184	2	BA35454
60	15	75.0	187	2	BA34778
61	15	75.0	194	1	AV054904
62	15	75.0	209	2	BA35996
63	15	75.0	218	4	BG655143
64	15	75.0	226	4	BA35445
65	15	75.0	227	2	BA35279
66	15	75.0	227	2	BA36610
67	15	75.0	228	1	AV061196
68	15	75.0	228	1	BA35171
69	15	75.0	228	2	BA35747
70	15	75.0	228	2	BA36012
71	15	75.0	228	2	BA35270
72	15	75.0	229	2	BA35411
73	15	75.0	229	2	BA35718
74	15	75.0	229	2	BA36377
75	15	75.0	230	2	BA34373
76	15	75.0	230	2	BA35180
77	15	75.0	230	2	BA35180
78	15	75.0	230	2	BA35493
79	15	75.0	231	1	AV060575
80	15	75.0	231	2	BA35274
81	15	75.0	231	2	BA35381
82	15	75.0	231	2	BA35513
83	15	75.0	231	2	BA35571
84	15	75.0	231	2	BA35580
85	15	75.0	232	2	BA35510
86	15	75.0	232	2	BA35337
87	15	75.0	232	2	BA35466
88	15	75.0	232	2	BA36903
89	15	75.0	236	2	BA35628
90	15	75.0	250	2	BA35696
91	15	75.0	251	2	BA35267
92	15	75.0	251	2	BA35935
93	15	75.0	251	2	BA36878
94	15	75.0	252	2	BA35143
95	15	75.0	252	2	BA35554
96	15	75.0	252	2	BA36273
97	15	75.0	253	2	BA34960

AZ982290	2M0263U04
BM704665	UI-E-C11-
CB177816	1821C01.X
BJ567993	BU567993
CB634129	OS1EB13H
BX254803	BX254803
AZ384762	1M0142124
BI113486	602899395
CN305165	170004250
BO829603	1L61n2145
CB840482	M1SE-0781
BX918106	BX918106
CB841003	M1SE-1452
AG295673	MUS MUSCU
BI221782	602937040
AZ342561	1M0075L12
AG422832	MUS MUSCU
BG921699	602825520
BE572798	601327831
BI869712	601393493
CO397391	ACENCCURT
BH096352	RPT1-24-3
BF123606	601760528
CL069801	CH216-117
BF344853	602014145
BO950615	ACENCCURT
CD838172	RFC02_110D
AA795871	VF284F01.T
BX511574	BX511574
CF652072	32-L02036
BP015388	BP015388
BA35135	BA35135
AW616050	ESTR296807
BA35118	BA35118
BA35454	BA35454
BA34778	BA34778
AV054904	AV054904
BA35996	BA35996
BG655143	1D25603.X
BA35445	BA35445
BA35279	BA35279
BA36610	BA36610
AV061196	AV061196
BA35171	BA35171
BA35747	BA35747
BA36012	BA36012
BA35270	BA35270
BA35411	BA35411
BA35718	BA35718
BA36377	BA36377
BA34373	BA34373
BA35180	BA35180
BA35493	BA35493
AV060575	AV060575
BA35274	BA35274
BA35381	BA35381
BA35513	BA35513
BA35571	BA35571
BA35580	BA35580
BA35510	BA35510
BA35337	BA35337
BA35466	BA35466
BA36903	BA36903
BA35628	BA35628
BA35696	BA35696
BA35267	BA35267
BA35935	BA35935
BA36878	BA36878
BA35143	BA35143
BA35554	BA35554
BA36273	BA36273
BA34960	BA34960

C 98 15 75.0 253 2 BB435272
C 99 15 75.0 253 2 BB435379
C 100 15 75.0 253 2 BB436103

ALIGNMENTS

RESULT 1
B1508715/c 613 bp mRNA linear EST 08-APR-2002
LOCUS B170011A10G03.5 Bee Brain Normalized/Subtracted Library, B17 Apis
DEFINITION mellifera cDNA clone B170011A10G03 5', mRNA sequence.

ACCESSION B1508715 GI:15359089
VERSION B1508715
KEYWORDS EST.
SOURCE Apis mellifera (honey bee)
ORGANISM Apis mellifera

REFERENCE Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.
AUTHORS Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L.,
PARTIALS, D., Robertson, H.M., Soares, B. and Robinson, G.E.
TITLE Annotated expressed sequence tags and cDNA microarrays for studies
of brain and behavior in the honey bee

JOURNAL Genome Res. 12 (4), 555-566 (2002)
MEDLINE 21929762
PUBMED 11932240

COMMENT Contact: Gene E. Robinson
Department of Entomology
University of Illinois
505 S. Goodwin Ave., Urbana, IL 61801, USA
Tel. 217 265 0308
Fax: 217 244 3499

Email: gene@life.uiuc.edu
This research was funded by the University of Illinois Critical
Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation
Award in Functional Genomics to G.E. Robinson and an NSF
Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.

PCR Primers
FORWARD: TAATGACATCTACTATGCG
BACKWARD: ATTAACTCTCCTAAG
Plate: B170011A10 row: G column: 03
Seq primer: AGCGATACATTCACACAGGA
High quality sequence stop: 613.

FEATURES

source

1..613
/organism="Apis mellifera"
/mol_type="mRNA"
/strain="mixed strains of European bees, predominantly
A.m. ligustica"
/db_xref="taxon:7460"
/clone="B170011A10G03"
/sex="female"
/tissue_type="brain"
/dev_stage="adult worker honey bee"
/lab_host="DH10B"
/clone_lib="Bee Brain Normalized/Subtracted Library, B17"
/note="Organ: brain; Vector: pRT73-Pac; Site 1: EcoRI;
Site 2: NotI; This B17 cDNA library was generated by
subtraction of the B16 library with 4000 previously
sequenced clones. The B16 library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806. RNA was
prepared from dissected brains of adult worker bees of
various ages and various behavioral groups."

ORIGIN

Query Match 95.0%; Score 19; DB 4; Length 613;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGGAGAGAGTCTCTTAT 20
|||||
Db 573 AGGAGAGAGTCTCTTAT 555

RESULT 2
BB434970/c 268 bp mRNA linear EST 01-AUG-2000
LOCUS BB434970 RIKEN full-length enriched, adult pancreas islet cells Mus
DEFINITION musculus cDNA clone C820010E15 3', mRNA sequence.

ACCESSION BB434970
VERSION BB434970.1 GI:9274697
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.
1 (bases 1 to 268)

REFERENCE Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
AUTHORS Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirokane, T., Horii, F., Ishii, Y., Ichikawa, J., Ichikawa, T., Itoh, M.,
Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,
Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Sugahara, Y., Suzuki, H., Tagawa, A.,
Takanashi, F., Tomimaga, N., Toyota, T., Tsunoda, Y., Watanabe, A.,
Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A.,
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.

RIKEN Mouse ESTs (Kono, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-reesgsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermostabilization and thermoinactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitanai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)

Please visit our web site (http://genome.riken.go.jp) for
further details.

FEATURES

source

1..268
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="C820010E15"
/tissue_type="pancreas"
/cell_type="Islet cells"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult pancreas
islet cells"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:28:49 ; Search time 34.4144 Seconds
(without alignments)
3138.794 Million cell updates/sec

Title: US-10-088-666-74

Perfect score: 20
Sequence: 1 gagggagaagatctctat 20

Scoring table: OLIGO-NUC , Gapex 60.0
Gapop 60.0 ,

Searched: 3625171 seqs, 2700493622 residues

Word size: 10

Total number of hits satisfying chosen parameters: 95006

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCRT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCRTS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	90.0	465	US-10-425-115-45008	Sequence 45008, A
2	17	85.0	127767	US-10-322-281-797	Sequence 797, App
3	16	80.0	688	US-10-027-632-102170	Sequence 102170, A
4	16	80.0	688	US-10-027-632-102171	Sequence 102171, A
5	16	80.0	688	US-10-027-632-102171	Sequence 102171, A
6	16	80.0	688	US-10-027-632-102171	Sequence 102171, A
7	16	80.0	2022	US-10-437-963-57769	Sequence 57769, A
8	16	80.0	6109	US-09-795-061-1	Sequence 1, Appli
9	16	80.0	6200	US-09-795-061-3	Sequence 3, Appli
10	16	80.0	6337	US-10-369-493-25224	Sequence 25224, A
11	16	80.0	218155	US-10-087-192-400	Sequence 400, App
12	15	75.0	255	US-10-149-736-7	Sequence 7, Appli

13	15	75.0	448	US-10-425-114-12489	Sequence 12489, A
14	15	75.0	448	US-10-425-115-140409	Sequence 140409, A
15	15	75.0	539	US-09-814-353-3404	Sequence 3404, Ap
16	15	75.0	539	US-09-814-353-9722	Sequence 9722, Ap
17	15	75.0	571	US-10-219-220-247	Sequence 247, App
18	15	75.0	605	US-09-938-842A-4589	Sequence 4589, Ap
19	15	75.0	605	US-09-938-842A-4589	Sequence 4589, Ap
20	15	75.0	1166	US-10-219-220-139	Sequence 139, App
21	15	75.0	1284	US-10-437-963-52109	Sequence 52109, A
22	15	75.0	1284	US-10-437-963-52109	Sequence 52109, A
23	15	75.0	1323	US-10-424-599-43260	Sequence 43260, A
24	15	75.0	1340	US-09-845-416-11	Sequence 11, Appli
25	15	75.0	1667	US-09-845-416-7	Sequence 7, Appli
26	15	75.0	1991	US-09-845-416-3	Sequence 3, Appli
27	15	75.0	2046	US-10-367-094-8	Sequence 8, Appli
28	15	75.0	3446	US-09-845-416-14	Sequence 14, Appli
29	15	75.0	3510	US-09-845-416-12	Sequence 12, Appli
30	15	75.0	3531	US-09-845-416-10	Sequence 10, Appli
31	15	75.0	3858	US-09-845-416-9	Sequence 9, Appli
32	15	75.0	3939	US-09-845-416-6	Sequence 6, Appli
33	15	75.0	4182	US-09-845-416-2	Sequence 2, Appli
34	15	75.0	4414	US-09-845-416-32	Sequence 32, Appli
35	15	75.0	4476	US-09-845-416-31	Sequence 31, Appli
36	15	75.0	4498	US-09-845-416-30	Sequence 30, Appli
37	15	75.0	4825	US-09-845-416-29	Sequence 29, Appli
38	15	75.0	4848	US-09-845-416-35	Sequence 35, Appli
39	15	75.0	4966	US-09-845-416-28	Sequence 28, Appli
40	15	75.0	4990	US-09-845-416-34	Sequence 34, Appli
41	15	75.0	5060	US-09-845-416-36	Sequence 36, Appli
42	15	75.0	5149	US-09-845-416-27	Sequence 27, Appli
43	15	75.0	5339	US-10-149-736-40	Sequence 40, Appli
44	15	75.0	5417	US-10-149-736-39	Sequence 39, Appli
45	15	75.0	5462	US-10-149-736-41	Sequence 41, Appli
46	15	75.0	8689	US-10-149-736-42	Sequence 42, Appli
47	15	75.0	11058	US-09-845-416-1	Sequence 1, Appli
48	15	75.0	11443	US-09-845-416-4	Sequence 4, Appli
49	15	75.0	12057	US-10-149-736-47	Sequence 47, Appli
50	15	75.0	13957	US-09-782-784-22	Sequence 22, Appli
51	15	75.0	13957	US-09-880-107-2284	Sequence 2284, Ap
52	15	75.0	13957	US-10-149-736-1	Sequence 1, Appli
53	15	75.0	14069	US-10-172-118-434	Sequence 434, App
54	15	75.0	14069	US-10-342-887-434	Sequence 434, App
55	15	75.0	14082	US-10-341-434-108	Sequence 108, App
56	15	75.0	14082	US-10-172-118-981	Sequence 981, App
57	15	75.0	14082	US-10-342-887-981	Sequence 981, App
58	15	75.0	66412	US-10-417-375-157	Sequence 157, App
59	15	75.0	108845	US-10-167-094-7	Sequence 7, Appli
60	15	75.0	175561	US-10-017-721-3	Sequence 3, Appli
61	15	75.0	175561	US-10-235-192A-48	Sequence 48, Appli
62	15	75.0	123025	US-10-289-762-1	Sequence 1, Appli
63	15	75.0	25	US-10-098-263B-16278	Sequence 16278, A
64	14	70.0	167	US-10-424-599-5970	Sequence 5970, Ap
65	14	70.0	204	US-10-674-124A-23903	Sequence 23903, A
66	14	70.0	247	US-10-674-124A-24560	Sequence 24560, A
67	14	70.0	257	US-10-219-220-138	Sequence 138, App
68	14	70.0	276	US-09-284-093B-2900	Sequence 2900, Ap
69	14	70.0	324	US-10-243-552-252	Sequence 252, App
70	14	70.0	333	US-10-260-238-5358	Sequence 5358, Ap
71	14	70.0	360	US-10-674-124A-15196	Sequence 15196, A
72	14	70.0	373	US-10-437-963-39058	Sequence 39058, A
73	14	70.0	371	US-10-425-115-173460	Sequence 173460, A
74	14	70.0	396	US-10-424-599-136954	Sequence 136954, A
75	14	70.0	426	US-10-027-632-86524	Sequence 86524, A
76	14	70.0	426	US-10-027-632-86525	Sequence 86525, A
77	14	70.0	426	US-10-027-632-86526	Sequence 86526, A
78	14	70.0	426	US-10-027-632-316407	Sequence 316407, A
79	14	70.0	426	US-10-027-632-316408	Sequence 316408, A
80	14	70.0	426	US-10-027-632-86524	Sequence 86524, A
81	14	70.0	426	US-10-027-632-86525	Sequence 86525, A
82	14	70.0	426	US-10-027-632-86526	Sequence 86526, A
83	14	70.0	426	US-10-027-632-316406	Sequence 316406, A
84	14	70.0	426	US-10-027-632-316407	Sequence 316407, A
85	14	70.0	426	US-10-027-632-316407	Sequence 316407, A

C 86 14 70.0 426 15 US-10-027-632-316408 Sequence 316408, A
C 87 14 70.0 436 16 US-10-424-599-79234 Sequence 79234, A
C 88 14 70.0 498 9 US-09-783-590-6021 Sequence 6021, Ap
C 89 14 70.0 544 16 US-10-660-226-15 Sequence 15, Ap1
90 14 70.0 569 13 US-10-044-739-1239 Sequence 1239, Ap
C 91 14 70.0 575 15 US-10-029-386-7633 Sequence 7633, Ap
C 92 14 70.0 579 15 US-10-084-817-82 Sequence 82, Ap1
C 93 14 70.0 599 13 US-09-734-569-129 Sequence 129, Ap1
C 94 14 70.0 599 13 US-10-027-632-63516 Sequence 63516, A
C 95 14 70.0 599 15 US-10-027-632-63516 Sequence 63516, A
C 96 14 70.0 611 13 US-10-027-632-321890 Sequence 221890, A
C 97 14 70.0 611 13 US-10-027-632-297152 Sequence 297152, A
C 98 14 70.0 611 15 US-10-027-632-221890 Sequence 221890, A
C 99 14 70.0 611 15 US-10-027-632-297152 Sequence 297152, A
100 14 70.0 615 16 US-10-282-122A-18935 Sequence 18935, A

ALIGNMENTS

RESULT 1
US-10-425-115-45008
; Sequence 45008, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 45008
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_141051C.1
US-10-425-115-45008

Query Match 90.0%; Score 18; DB 18; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGGAAGAAGTCTCTTAT 20
DB 439 GGGAAGAAGTCTCTTAT 456

RESULT 2
US-10-322-281-797
; Sequence 797, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 797
; LENGTH: 127767
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(127767)
; OTHER INFORMATION: n = A, T, C or G

US-10-322-281-797
Query Match 85.0%; Score 17; DB 17; Length 127767;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGAAGAAGTCTCTTAT 20
DB 5129 GGGAAGAAGTCTCTTAT 5145

RESULT 3
US-10-027-632-102170/c
; Sequence 102170, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102170
; LENGTH: 688
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-102170

Query Match 80.0%; Score 16; DB 13; Length 688;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGGGAAGAAGTCTCT 17
DB 631 AGGGAAGAAGTCTCT 616

RESULT 4
US-10-027-632-102171/c
; Sequence 102171, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:18:54 ; Search time 8.52853 Seconds
(without alignments)
1666.850 Million cell updates/sec

Title: US-10-088-666-74

Perfect score: 20

Sequence: 1 gagggagaagatctctat 20

Scoring table: OIIGO-NUC;
Gapop 60.0 , Gapext 60.0

Searched: 824507 seqs, 355394441 residues

Word size: 10

Total number of hits satisfying chosen parameters: 9618

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/ECTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	80.0	6109	4	US-09-795-061-1
2	16	80.0	6200	4	US-09-795-061-3
3	15	75.0	1166	4	US-09-325-932A-139
4	15	75.0	5952	4	US-09-687-875A-1
5	15	75.0	13977	4	US-09-484-970B-60
6	15	75.0	1230025	4	US-09-198-452A-1
7	14	70.0	257	4	US-09-325-932A-138
8	14	70.0	479	4	US-09-621-976-893
9	14	70.0	858	4	US-09-107-532A-659
10	14	70.0	1122	4	US-09-252-991A-12729
11	14	70.0	1152	4	US-09-252-991A-12886
12	14	70.0	1269	4	US-09-625-198-11
13	14	70.0	1683	3	US-09-192-545-1
14	14	70.0	1743	4	US-09-252-991A-13058
15	14	70.0	2452	4	US-09-799-451-786
16	14	70.0	3001	4	US-09-539-333D-156
17	14	70.0	3880	4	US-09-221-017B-1071
18	14	70.0	4053	4	US-09-620-312D-156
19	14	70.0	4519	3	US-08-976-289-6
20	14	70.0	4519	4	US-09-956-004-6
21	14	70.0	4948	4	US-09-562-702A-23
22	14	70.0	4948	4	US-09-561-818A-23
23	14	70.0	4972	4	US-09-562-702A-27
24	14	70.0	5306	4	US-09-562-702A-21
25	14	70.0	5306	4	US-09-561-818A-21
26	14	70.0	5306	4	US-09-562-702A-25
27	14	70.0	40123	4	US-08-311-731A-137

C 28	13	65.0	120	4	US-09-270-767-30650	Sequence 30650, A
C 29	13	65.0	196	4	US-09-513-999C-27882	Sequence 27882, A
C 30	13	65.0	246	4	US-09-513-999C-36601	Sequence 36601, A
C 31	13	65.0	258	4	US-09-513-999C-10393	Sequence 10393, A
C 32	13	65.0	266	4	US-09-270-767-2288	Sequence 2288, A
C 33	13	65.0	266	4	US-09-270-767-17570	Sequence 17570, A
C 34	13	65.0	288	4	US-09-248-796A-3763	Sequence 3763, A
C 35	13	65.0	361	4	US-09-270-767-9402	Sequence 9402, A
C 36	13	65.0	361	4	US-09-270-767-24684	Sequence 24684, A
C 37	13	65.0	412	4	US-09-513-999C-562	Sequence 562, A
C 38	13	65.0	441	3	US-09-222-575-114	Sequence 114, A
C 39	13	65.0	441	4	US-09-389-681-114	Sequence 114, A
C 40	13	65.0	441	4	US-09-620-405B-114	Sequence 114, A
C 41	13	65.0	441	4	US-09-339-338-114	Sequence 114, A
C 42	13	65.0	441	4	US-09-433-826B-114	Sequence 114, A
C 43	13	65.0	441	4	US-09-604-876-114	Sequence 114, A
C 44	13	65.0	441	4	US-09-285-480-114	Sequence 114, A
C 45	13	65.0	441	4	US-09-834-759-114	Sequence 114, A
C 46	13	65.0	441	4	US-09-590-751A-114	Sequence 114, A
C 47	13	65.0	442	4	US-09-270-767-10431	Sequence 10431, A
C 48	13	65.0	470	2	US-08-967-101-77	Sequence 77, A
C 49	13	65.0	470	2	US-08-592-541-77	Sequence 77, A
C 50	13	65.0	470	3	US-09-124-698-77	Sequence 77, A
C 51	13	65.0	470	3	US-09-127-480-77	Sequence 77, A
C 52	13	65.0	470	3	US-08-496-841C-77	Sequence 77, A
C 53	13	65.0	470	3	US-09-124-523-77	Sequence 77, A
C 54	13	65.0	470	4	US-09-636-796A-77	Sequence 77, A
C 55	13	65.0	470	4	US-08-431-048F-77	Sequence 77, A
C 56	13	65.0	472	4	US-09-621-976-12821	Sequence 12821, A
C 57	13	65.0	493	4	US-09-621-976-12217	Sequence 12217, A
C 58	13	65.0	631	4	US-09-270-767-10321	Sequence 10321, A
C 59	13	65.0	766	4	US-09-328-352-2137	Sequence 2137, A
C 60	13	65.0	840	2	US-08-924-759-13	Sequence 13, A
C 61	13	65.0	840	3	US-09-248-335-13	Sequence 13, A
C 62	13	65.0	844	3	US-08-998-416-292	Sequence 292, A
C 63	13	65.0	877	4	US-09-270-767-14474	Sequence 14474, A
C 64	13	65.0	885	4	US-09-583-110-1634	Sequence 1634, A
C 65	13	65.0	948	4	US-09-270-767-11984	Sequence 11984, A
C 66	13	65.0	971	3	US-09-248-335-65	Sequence 65, A
C 67	13	65.0	1001	4	US-09-641-638-312	Sequence 312, A
C 68	13	65.0	1001	4	US-10-170-097-312	Sequence 312, A
C 69	13	65.0	1008	4	US-09-940-244-348	Sequence 348, A
C 70	13	65.0	1020	4	US-09-107-532A-2766	Sequence 2766, A
C 71	13	65.0	1100	4	US-09-270-767-14335	Sequence 14335, A
C 72	13	65.0	1200	4	US-09-543-681A-761	Sequence 761, A
C 73	13	65.0	1221	4	US-09-799-451-421	Sequence 421, A
C 74	13	65.0	1332	4	US-09-976-594-310	Sequence 310, A
C 75	13	65.0	1338	4	US-09-543-681A-2715	Sequence 2715, A
C 76	13	65.0	1367	4	US-09-167-206-5	Sequence 5, A
C 77	13	65.0	1559	3	US-08-747-574-2	Sequence 2, A
C 78	13	65.0	1827	4	US-09-252-991A-1646	Sequence 1466, A
C 79	13	65.0	1836	3	US-08-795-430-10	Sequence 10, A
C 80	13	65.0	1836	3	US-09-355-700-10	Sequence 10, A
C 81	13	65.0	1836	4	US-08-601-132-40	Sequence 40, A
C 82	13	65.0	1836	4	US-08-671-573B-40	Sequence 40, A
C 83	13	65.0	1836	4	US-09-631-092B-40	Sequence 40, A
C 84	13	65.0	1842	4	US-09-81-760B-33	Sequence 33, A
C 85	13	65.0	1836	3	US-09-332-200-30	Sequence 30, A
C 86	13	65.0	1836	3	US-09-232-197-30	Sequence 30, A
C 87	13	65.0	1836	3	US-09-232-197-30	Sequence 30, A
C 88	13	65.0	1836	3	US-09-332-197-31	Sequence 31, A
C 89	13	65.0	1836	3	US-09-332-201-30	Sequence 30, A
C 90	13	65.0	1836	3	US-09-332-201-31	Sequence 31, A
C 91	13	65.0	1836	4	US-09-332-195-30	Sequence 30, A
C 92	13	65.0	1836	4	US-09-332-195-31	Sequence 31, A
C 93	13	65.0	1908	4	US-09-328-352-1340	Sequence 1340, A
C 94	13	65.0	1927	1	US-08-336-583-1	Sequence 1, A
C 95	13	65.0	2193	5	US-09-529-1379B-1	Sequence 1, A
C 96	13	65.0	2193	4	US-09-529-812D-688	Sequence 688, A
C 97	13	65.0	2191	3	US-09-599-812B-25	Sequence 25, A
C 98	13	65.0	2191	3	US-09-298-731-25	Sequence 25, A
C 99	13	65.0	2191	4	US-09-350-614-25	Sequence 25, A
C 100	13	65.0	2301	3	US-09-232-191-8	Sequence 8, A

ALIGNMENTS

RESULT 1

US-09-795-061-1
; Sequence 1, Application US/09795061
; Patent No. 6759528
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Imamura, Yasutada
; TITLE OF INVENTION: Pro-Alpha 3 (V) Collagen Genes
; FILE REFERENCE: 960296.96781
; CURRENT APPLICATION NUMBER: US/09/795,061
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6109
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (82)..(5298)
US-09-795-061-1

Query Match 80.0%; Score 16; DB 4; Length 6109;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGAGAGTCTTC 16
DB 4907 GAGGAGAGAGTCTTC 4922

RESULT 2
US-09-795-061-3
; Sequence 3, Application US/09795061
; Patent No. 6759528
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Imamura, Yasutada
; TITLE OF INVENTION: Pro-Alpha 3 (V) Collagen Genes
; FILE REFERENCE: 960296.96781
; CURRENT APPLICATION NUMBER: US/09/795,061
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(5321)
US-09-795-061-3

Query Match 80.0%; Score 16; DB 4; Length 6200;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGAGAGTCTTC 16
DB 4927 GAGGAGAGAGTCTTC 4942

RESULT 3
US-09-325-932A-139/c
; Sequence 139, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry

; APPLICANT: Iasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 139
; LENGTH: 1166
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-139

Query Match 75.0%; Score 15; DB 4; Length 1166;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAGAGTCTCTTAT 20
DB 1001 AAGAGTCTCTTAT 987

RESULT 4
US-09-687-875A-1/c
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 75.0%; Score 15; DB 4; Length 5952;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGGAGAGAGTCTCT 17
DB 874 GGGAGAGAGTCTCT 860

RESULT 5
US-09-484-970B-60/c
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:59:59 ; Search time 38.7387 Seconds
(without alignments)
2710.166 Million cell updates/sec

Title: US-10-088-666-74

Perfect score: 20

Sequence: 1 gaggaagaagttctctat 20

Scoring table: OLIGO_NUC;
Gapop 60.0, Gapext 60.0

Searched: 413486 segs, 2624710521 residues

Word size: 10

Total number of hits satisfying chosen parameters: 86295

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1900s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	5	AA61639
2	20	100.0	267	5	AA61639 Lactobaci
3	16	80.0	337	6	ABN26443
4	16	80.0	995	10	ADK53632
5	16	80.0	995	10	ADK55905
6	16	80.0	6109	4	AAH78667
7	16	80.0	6200	4	AAH78668
8	16	80.0	6200	12	ADQ19840
9	16	80.0	6368	6	ABN59647
10	16	80.0	6617	12	ADQ23958
11	15	75.0	255	6	ABK81965
12	15	75.0	445	2	AAH91154
13	15	75.0	500	6	ABQ63274
14	15	75.0	535	3	AA52735
15	15	75.0	539	5	AD170662
16	15	75.0	539	5	AD176980
17	15	75.0	571	10	ADB94883
18	15	75.0	605	6	AB217184
19	15	75.0	661	2	AAH91127
20	15	75.0	702	2	AA592053
21	15	75.0	1034	2	AAQ06255

22	15	75.0	1034	2	AA773216	Act73216 Rat prepr
23	15	75.0	1034	2	AA220678	AA220678 Preproglu
24	15	75.0	1166	4	AA44834	AA44834 RPP5-Like
25	15	75.0	1166	10	ADB94775	ADB94775 Programme
26	15	75.0	1340	6	AA037233	AA037233 Human dys
27	15	75.0	1667	6	AA037235	AA037235 Human dys
28	15	75.0	1991	6	AA037231	AA037231 Human dys
29	15	75.0	2230	3	AA054329	AA054329 Arabidops
30	15	75.0	3446	3	AA248570	AA248570 A rod sho
31	15	75.0	3510	6	AA037242	AA037242 Human dys
32	15	75.0	3531	6	AA037238	AA037238 Human dys
33	15	75.0	3747	3	AA248566	AA248566 A rod sho
34	15	75.0	3858	6	AA037237	AA037237 Human dys
35	15	75.0	3999	6	AA037234	AA037234 Human dys
36	15	75.0	4075	3	AA248569	AA248569 A rod sho
37	15	75.0	4182	6	AA037230	AA037230 Human dys
38	15	75.0	4402	3	AA248568	AA248568 A rod sho
39	15	75.0	4402	3	AA248567	AA248567 A rod sho
40	15	75.0	4414	6	AA037260	AA037260 Adeno-ass
41	15	75.0	4476	6	AA037259	AA037259 Adeno-ass
42	15	75.0	4476	6	AA037258	AA037258 Adeno-ass
43	15	75.0	4825	6	AA037257	AA037257 Adeno-ass
44	15	75.0	4848	6	AA037263	AA037263 Adeno-ass
45	15	75.0	4966	6	AA037256	AA037256 Adeno-ass
46	15	75.0	4990	6	AA037262	AA037262 Adeno-ass
47	15	75.0	5060	6	AA037264	AA037264 Adeno-ass
48	15	75.0	5149	6	AA037255	AA037255 Adeno-ass
49	15	75.0	5268	6	AB017929	AB017929 Human ner
50	15	75.0	5339	6	ABK81998	ABK81998 DNA encod
51	15	75.0	5417	6	ABK81997	ABK81997 DNA encod
52	15	75.0	5462	6	ABK81999	ABK81999 DNA encod
53	15	75.0	5952	5	AA06794	AA06794 Human ner
54	15	75.0	6798	5	AB015941	AB015941 Human dys
55	15	75.0	6798	6	ABN95786	ABN95786 Gene #228
56	15	75.0	13957	6	ABN95786	ABN95786 Gene #228
57	15	75.0	13957	6	ABN95786	ABN95786 Gene #228
58	15	75.0	11089	6	AA037229	AA037229 Human dys
59	15	75.0	11241	6	ABK82005	ABK82005 cDNA encod
60	15	75.0	11443	6	ABK82002	ABK82002 cDNA encod
61	15	75.0	12923	1	AA090338	AA090338 Sequence
62	15	75.0	13957	6	ABK81959	ABK81959 cDNA encod
63	15	75.0	13957	6	ABT10904	ABT10904 Human bre
64	15	75.0	13957	6	ABN95786	ABN95786 Gene #228
65	15	75.0	13977	6	ABN95786	ABN95786 Gene #228
66	15	75.0	13977	6	ABN95786	ABN95786 Gene #228
67	15	75.0	13977	6	ABN95786	ABN95786 Gene #228
68	15	75.0	14069	12	ADN04003	ADN04003 Antipso
69	15	75.0	21082	5	AB017928	AB017928 Human ner
70	15	75.0	21082	5	AB017927	AB017927 Human ner
71	15	75.0	21989	5	AB015942	AB015942 Human ner
72	15	75.0	21989	5	AB015942	AB015942 Human ner
73	15	75.0	21989	5	AB015942	AB015942 Human ner
74	15	75.0	21989	5	AB015942	AB015942 Human ner
75	15	75.0	21989	5	AB015942	AB015942 Human ner
76	15	75.0	21989	5	AB015942	AB015942 Human ner
77	15	75.0	21989	5	AB015942	AB015942 Human ner
78	15	75.0	21989	5	AB015942	AB015942 Human ner
79	15	75.0	21989	5	AB015942	AB015942 Human ner
80	15	75.0	21989	5	AB015942	AB015942 Human ner
81	15	75.0	21989	5	AB015942	AB015942 Human ner
82	15	75.0	21989	5	AB015942	AB015942 Human ner
83	15	75.0	21989	5	AB015942	AB015942 Human ner
84	15	75.0	21989	5	AB015942	AB015942 Human ner
85	15	75.0	21989	5	AB015942	AB015942 Human ner
86	15	75.0	21989	5	AB015942	AB015942 Human ner
87	15	75.0	21989	5	AB015942	AB015942 Human ner
88	15	75.0	21989	5	AB015942	AB015942 Human ner
89	15	75.0	21989	5	AB015942	AB015942 Human ner
90	15	75.0	21989	5	AB015942	AB015942 Human ner
91	15	75.0	21989	5	AB015942	AB015942 Human ner
92	15	75.0	21989	5	AB015942	AB015942 Human ner
93	15	75.0	21989	5	AB015942	AB015942 Human ner
94	15	75.0	21989	5	AB015942	AB015942 Human ner

95 14 70.0 831 5 AAS88057 DNA encod
96 14 70.0 831 5 AAS64236 DNA encod
97 14 70.0 856 6 ABQ44964 Oligonuc1
98 14 70.0 856 6 ABQ44965 Oligonuc1
99 14 70.0 858 10 ADC91032 E. faeciu
100 14 70.0 1122 11 ABD14125 Pseudomon

ALIGNMENTS

RESULT 1

AA61639

ID AAF61639 standard; DNA; 20 BP.

AA61639;

02-JUL-2001 (first entry)

Lactobacillus brevis 23S rRNA/5S rRNA specific probe SEQ ID 74.

23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

Lactobacillus brevis.

DE19945964-A1.

05-APR-2001.

24-SEP-1999; 99DE-01045964.

24-SEP-1999; 99DE-01045964.

(BIOT-) BIOTECON DIAGNOSTICS GMBH.

Fandke M, Gasch A, Berghof K;

WPI; 2001-246136/26.

Detecting contaminating microorganisms in brewing, by nucleic acid

amplification and hybridization, either non-specific or genus- or species

specific.

Claim 9(i); Page 17; 48pp; German.

This invention describes a novel method for detecting microorganisms (A) of importance in brewing which comprises treating a sample with at least two primers (P1) that hybridize to a consensus region in the nucleic acid of (A), at least part of the microbial nucleic acid is amplified, the amplicon is treated with at least one probe (P2) that hybridizes specifically with a sequence common to all (A) or specific for one or more families, genera or species, and any formation of hybrids is detected. The method is used to detect, identify and/or characterize microorganisms in beer or brewing materials, particularly for detecting contamination. The method may detect the entire range of contaminating microbes, either as a general test for contamination or as a test specific for particular genera or (sub)species. It is quicker than known microbiological methods, and can detect several organisms in the same sample, including organisms not presently recognized as contaminants. The method provides an early indication of contamination and can be automated for high throughput analysis

Sequence 20 BP; 6 A; 2 C; 6 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 5; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAGGGAAGAAGTCTCTTAT 20

1 GAGGGAAGAAGTCTCTTAT 20

AA61566
ID AAF61566 standard; DNA; 267 BP.

AA61566;

02-JUL-2001 (first entry)

Lactobacillus brevis 23S rRNA-spacer-5S rRNA DNA fragment.

23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

Lactobacillus brevis.

DE19945964-A1.

05-APR-2001.

24-SEP-1999; 99DE-01045964.

24-SEP-1999; 99DE-01045964.

(BIOT-) BIOTECON DIAGNOSTICS GMBH.

Fandke M, Gasch A, Berghof K;

WPI; 2001-246136/26.

Detecting contaminating microorganisms in brewing, by nucleic acid amplification and hybridization, either non-specific or genus- or species specific.

Claim 9(i); Page 9; 48pp; German.

This invention describes a novel method for detecting microorganisms (A) of importance in brewing which comprises treating a sample with at least two primers (P1) that hybridize to a consensus region in the nucleic acid of (A), at least part of the microbial nucleic acid is amplified, the amplicon is treated with at least one probe (P2) that hybridizes specifically with a sequence common to all (A) or specific for one or more families, genera or species, and any formation of hybrids is detected. The method is used to detect, identify and/or characterize microorganisms in beer or brewing materials, particularly for detecting contamination. The method may detect the entire range of contaminating microbes, either as a general test for contamination or as a test specific for particular genera or (sub)species. It is quicker than known microbiological methods, and can detect several organisms in the same sample, including organisms not presently recognized as contaminants. The method provides an early indication of contamination and can be automated for high throughput analysis

Sequence 267 BP; 77 A; 51 C; 77 G; 62 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 5; Length 267;

Best Local Similarity 100.0%; Pred. No. 0.29;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAGGGAAGAAGTCTCTTAT 20

154 GAGGGAAGAAGTCTCTTAT 173

RESULT 3

ABN26443

ID ABN26443 standard; cDNA; 337 BP.

ABN26443;

24-JUN-2002 (first entry)

Human ORFX polymynucleotide sequence SEQ ID NO:21363.

Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 12:43:29 / Search time 156.156 Seconds

(without alignments)
6056.719 Million cell updates/sec

Title: US-10-088-666-74

Perfect score: 20

Sequence: 1 gagggaagaagtcctctat 20

Scoring table: OLIGO-NUC

Gapop 60.0, Gapext 60.0

Searched: 4526729 segs, 23644849745 residues

Word-size: 10

Total number of hits satisfying chosen parameters: 353379

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database:

GenEmbl:
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_rts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	AX105802	AX105802 Sequence
2	20	100.0	267	AX105729	AX105729 Sequence
3	18	90.0	229813	AC117613	AC117613 Mus muscu
4	17	85.0	101707	AC150157	AC150157 Gallus ga
5	17	85.0	128809	AC021648	AC021648 Homo sapi
6	17	85.0	172825	AC020614	AC020614 Homo sapi
7	17	85.0	187221	AC117229	AC117229 Mus muscu
8	17	85.0	192734	AL928798	AL928798 Mouse DNA
9	17	85.0	192944	AC146302	AC146302 Callicebu
10	17	85.0	227515	AC150065	AC150065 Gallus ga
11	17	85.0	234651	AC115456	AC115456 Rattus no
12	16	80.0	337	CQ455863	CQ455863 Sequence
13	16	80.0	1473	AY039654	AY039654 Danto rer
14	16	80.0	1907	CQ723784	CQ723784 Sequence
15	16	80.0	2638	RABMBB4	RABMBB4
16	16	80.0	6076	AF272661	AF272661 rabbit emb
17	16	80.0	6109	AX239611	AX239611 Rattus no
18	16	80.0	6109	AF176645	AF176645 Mus muscu
19	16	80.0	6200	AX239613	AX239613 Sequence

20	16	80.0	6200	9	AF177941	AF177941 Homo sapi
21	16	80.0	6368	6	AX405643	AX405643 Sequence
22	16	80.0	7374	8	SCCDC	X70151 S. cerevisi
23	16	80.0	44594	4	RABRGLB	M18818 Rabbit beta
24	16	80.0	72872	4	FL104	AF096370 Arabidops
25	16	80.0	76425	9	AC105346	AC105346 Homo sapi
26	16	80.0	105907	2	AC026429	AC026429 Homo sapi
27	16	80.0	110000	1	EX908798	Continuation (2 of
28	16	80.0	110000	2	AC116152	Continuation (3 of
29	16	80.0	110000	2	AC149068	AC149068 Mus muscu
30	16	80.0	112945	8	AC027683	AC027683 Homo sapi
31	16	80.0	135785	8	AC051634	AC051634 Oryza sat
32	16	80.0	141773	2	AC069348	AC069348 Homo sapi
33	16	80.0	141790	10	AL671759	AL671759 Mouse DNA
34	16	80.0	144207	2	AC115887	AC115887 Mus muscu
35	16	80.0	150872	8	AP003574	AP003574 Oryza sat
36	16	80.0	153616	2	AC136945	AC136945 Homo sapi
37	16	80.0	156054	9	HSDJ16705	AL121938 Human DNA
38	16	80.0	156791	9	HSDJ860P4	AL045594 Human DNA
39	16	80.0	158756	2	AC137793	AC137793 Homo sapi
40	16	80.0	163757	2	AC141448	AC141448 Homo sapi
41	16	80.0	165500	9	AC093784	AC093784 Homo sapi
42	16	80.0	167231	9	AC099571	AC099571 Homo sapi
43	16	80.0	169765	9	AC006075	AC006075 Homo sapi
44	16	80.0	170106	9	AC104371	AC104371 Homo sapi
45	16	80.0	176625	8	AP004992	AP004992 Oryza sat
46	16	80.0	177029	2	AC079079	AC079079 Homo sapi
47	16	80.0	178366	2	AC142217	AC142217 Rattus no
48	16	80.0	178397	2	AC131789	AC131789 Mus muscu
49	16	80.0	183067	9	AC040914	AC040914 Homo sapi
50	16	80.0	183583	9	AC068408	AC068408 Homo sapi
51	16	80.0	184909	9	AL356057	AL356057 Human DNA
52	16	80.0	189616	2	AC121093	AC121093 Mus muscu
53	16	80.0	189628	2	AC146318	AC146318 Gallus ga
54	16	80.0	198697	8	ATCRRIV4	AL161492 Arabidops
55	16	80.0	198788	8	ATCRRIV46	AL161546 Arabidops
56	16	80.0	199825	9	AC099570	AC099570 Homo sapi
57	16	80.0	200891	2	AC109815	AC109815 Mus muscu
58	16	80.0	202576	10	AC140931	AC140931 Mus muscu
59	16	80.0	207674	8	ATFCB8	297343 Arabidops
60	16	80.0	212437	2	AC139865	AC139865 Mus muscu
61	16	80.0	218632	2	AC138333	AC138333 Mus muscu
62	16	80.0	220313	2	AC145406	AC145406 Gallus ga
63	16	80.0	221478	2	AC079545	AC079545 Mus muscu
64	16	80.0	222159	2	AC106540	AC106540 Rattus no
65	16	80.0	229114	2	AC106554	AC106554 Rattus no
66	16	80.0	234182	2	AC111393	AC111393 Rattus no
67	16	80.0	236182	2	AC131024	AC131024 Rattus no
68	16	80.0	238147	3	CNS07EGF	AL590448 chromoso
69	16	80.0	245671	2	AC112762	AC112762 Rattus no
70	16	80.0	247462	2	AC134478	AC134478 Rattus no
71	16	80.0	258445	2	AC126131	AC126131 Rattus no
72	16	80.0	261665	2	AC095565	AC095565 Rattus no
73	16	80.0	267520	2	AC095489	AC095489 Rattus no
74	16	80.0	316613	8	AE011113	AE011113 Oryza sat
75	16	80.0	316613	1	AY434428	AY434428 S. cerevisi
76	16	75.0	173	1	SCCHR11	AY973702 S. cerevisi
77	15	75.0	255	6	AX538587	AX538587 Sequence
78	15	75.0	334	9	AF213409	AF213409 Homo sapi
79	15	75.0	385	1	AF078273	AF078273 Grassland
80	15	75.0	414	1	AF078284	AF078284 Grassland
81	15	75.0	428	1	AY214627	AY214627 Unculture
82	15	75.0	445	1	AF365745	AF365745 Unculture
83	15	75.0	445	6	AF64103	AF64103 Sequence
84	15	75.0	445	1	AY785039	AY785039 Sequence
85	15	75.0	468	1	UBA421881	UBA421881 unculture
86	15	75.0	500	6	AX474829	AX474829 Sequence
87	15	75.0	509	10	RATGLD6	K02813 Rat glucago
88	15	75.0	535	8	AY084318	AY084318 Arabidops
89	15	75.0	539	6	CQ396333	CQ396333 Sequence
90	15	75.0	539	6	CQ402651	CQ402651 Sequence
91	15	75.0	553	1	AF432751	AF432751 unculture
92	15	75.0	605	6	AX510294	AX510294 Sequence

93 15 75.0 661 6 A64076
 94 15 75.0 661 6 AX785012
 95 15 75.0 669 1 AY093424
 96 15 75.0 674 8 AF041386
 97 15 75.0 682 1 AB064709
 98 15 75.0 700 8 FV41446
 99 15 75.0 731 1 AY463178
 100 15 75.0 737 8 AF082300

ALIGNMENTS

RESULT 1
 AX105802
 LOCUS AX105802
 DEFINITION Sequence 74 from Patent WO0123605.
 ACCESSION AX105802
 VERSION AX105802.1 GI:13921815
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM artificial sequences.

REFERENCE
 AUTHORS Fandke, M., Gasch, A. and Bergshof, K.
 TITLE Method and nucleic acids for determining the presence of
 JOURNAL micro-organisms specific to the brewing process
 PATENT: WO 0123605-A 74 05-APR-2001;
 Biotechn Diagnostics GmbH (DE)
 FEATURES
 source Location/Qualifiers
 1..20
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="spezifische Sequenz fur Lactobacillus brevis"

ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGGAAGAAGTCTCTTAT 20
 1 GAGGGAAGAAGTCTCTTAT 20

RESULT 2
 AX105729
 LOCUS AX105729
 DEFINITION Sequence 1 from Patent WO0123605.
 ACCESSION AX105729
 VERSION AX105729.1 GI:13921742
 KEYWORDS
 SOURCE Lactobacillus brevis
 ORGANISM Lactobacillus brevis

REFERENCE
 AUTHORS Fandke, M., Gasch, A. and Bergshof, K.
 TITLE Method and nucleic acids for determining the presence of
 JOURNAL micro-organisms specific to the brewing process
 PATENT: WO 0123605-A 1 05-APR-2001;
 Biotechn Diagnostics GmbH (DE)
 FEATURES
 source Location/Qualifiers
 1..267
 /organism="Lactobacillus brevis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1580"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 267;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGGAAGAAGTCTCTTAT 20
 154 GAGGGAAGAAGTCTCTTAT 173

RESULT 3
 AC117613/c
 LOCUS AC117613
 DEFINITION 229813 bp DNA linear HTG 25-MAR-2004
 Mus musculus chromosome 6 clone RP23-106C18 map 6, *** SEQUENCING
 IN PROGRESS ***, 9 unordered pieces.

ACCESSION AC117613
 AC117613.6 GI:45736872
 VERSION
 HTG; HTGS PHASE1; HTGS_FULFILL; HTGS_ACTIVEFIN.
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 229813)
 REFERENCE
 AUTHORS Birren, B., Nussbaum, C. and Lander, E.
 TITLE Mus musculus chromosome 6, clone RP23-106C18
 JOURNAL Unpublished
 REFERENCE
 AUTHORS 2 (bases 1 to 229813)

Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, X., Dewar, K., Diaz, J., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Labrecque, K., Lamazeres, R.,
 Landers, T., Lehoczek, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Melgrim, J., Meneus, L.,
 Miho, T., Miho, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, C., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, W., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnuppback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggio, D., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE
 AUTHORS 3 (bases 1 to 229813)

Birren, B., Nussbaum, C., Lander, E., Aboueleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H. M., Barina, N., Bastien, V.,
 Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 Dearellano, K., Diaz, J., Dodge, S., Dooley, K., Dorris, L.,
 Galagan, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 Maclean, C., MacDonald, P., Major, J., Manning, J., Meneus, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Miho, V., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnuppback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, U.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (25-MAR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:16:19 ; Search time 144.52 Seconds
(without alignments)
6303.600 Million cell updates/sec

Title: US-10-088-666-73

Perfect score: 25

Sequence: 1 tcgagataatgataatcctag 25

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size: 10

Total number of hits satisfying chosen parameters: 675855

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Listing first 100 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	76.0	877	8	CG141515
2	19	76.0	888	9	CG965204
3	18	72.0	668	7	CR286805
4	18	72.0	738	7	CR286927
5	18	72.0	797	8	BZ192101
6	18	72.0	808	7	CK362822
7	18	72.0	848	8	AZ683885
8	18	72.0	898	8	BH133066
9	18	72.0	900	8	BH147215
10	18	72.0	903	8	BH136672
11	18	72.0	1227	9	CG751219
12	17	68.0	407	8	AZ176707
13	17	68.0	458	8	AZ177785
14	17	68.0	506	9	CG696733
15	17	68.0	512	8	AZ955016
16	17	68.0	513	8	AQ382746
17	17	68.0	513	9	CG758285
18	17	68.0	519	9	CG256503
19	17	68.0	534	8	AQ819623
20	17	68.0	535	8	BH829578
21	17	68.0	544	4	BI467380
22	17	68.0	546	8	CT583870
23	17	68.0	553	8	BH099534
24	17	68.0	574	8	AQ188728

25	17	68.0	579	9	CC846880
26	17	68.0	582	4	BT693244
27	17	68.0	584	8	AQ261596
28	17	68.0	591	8	BH766079
29	17	68.0	594	1	AU290058
30	17	68.0	607	8	BH820899
31	17	68.0	615	9	CC858148
32	17	68.0	627	8	CC858148
33	17	68.0	629	7	CK818278
34	17	68.0	671	9	CT551967
35	17	68.0	680	1	AU217527
36	17	68.0	695	4	BT702183
37	17	68.0	704	9	AG331381
38	17	68.0	724	6	CF264610
39	17	68.0	735	8	BZ847658
40	17	68.0	741	4	BT815333
41	17	68.0	745	8	CE770220
42	17	68.0	776	8	BH487731
43	17	68.0	781	8	AZ210199
44	17	68.0	786	7	CK026066
45	17	68.0	786	7	CC816891
46	17	68.0	875	6	CA474062
47	17	68.0	912	8	CC133823
48	17	68.0	943	8	AZ136157
49	17	68.0	1398	9	CL022024
50	16	64.0	198	9	BX285738
51	16	64.0	233	9	AL946816
52	16	64.0	257	1	A1626105
53	16	64.0	263	9	CR401435
54	16	64.0	280	4	BG211404
55	16	64.0	298	7	D72160
56	16	64.0	317	2	BB233201
57	16	64.0	328	7	CN924215
58	16	64.0	365	7	CN453602
59	16	64.0	367	8	AQ194703
60	16	64.0	380	9	BX221357
61	16	64.0	383	8	AQ410152
62	16	64.0	402	1	AU061419
63	16	64.0	415	4	BM599153
64	16	64.0	421	1	AJ672916
65	16	64.0	422	6	CB909797
66	16	64.0	429	6	CA865978
67	16	64.0	430	9	BX285737
68	16	64.0	434	6	CD450045
69	16	64.0	438	8	AQ829583
70	16	64.0	440	1	BJ435871
71	16	64.0	441	1	A1489028
72	16	64.0	460	8	AQ661945
73	16	64.0	464	2	BF003557
74	16	64.0	464	4	BM359958
75	16	64.0	469	4	AQ593197
76	16	64.0	480	6	CD842212
77	16	64.0	487	9	CE128067
78	16	64.0	493	6	CD195244
79	16	64.0	493	8	AQ155496
80	16	64.0	495	8	AQ245575
81	16	64.0	500	8	BH527994
82	16	64.0	500	8	BZ347068
83	16	64.0	500	8	BZ351646
84	16	64.0	502	6	AQ130690
85	16	64.0	502	6	CA841028
86	16	64.0	503	7	CK523454
87	16	64.0	516	9	BX237955
88	16	64.0	517	8	AZ925835
89	16	64.0	519	9	CNS00VM9
90	16	64.0	521	8	AZ927695
91	16	64.0	522	7	CO276626
92	16	64.0	522	8	B61330
93	16	64.0	534	8	AQ060625
94	16	64.0	534	8	AQ060625
95	16	64.0	547	8	AQ052705
96	16	64.0	549	4	AQ565565
97	16	64.0	551	4	BJ482831
98	16	64.0	551	1	AU262981

CC846880	NDL.129K1
BT693244	BT693244
AQ261596	CTBT1-E1-
BH766079	BH8AC359E
AU290058	AU290058
BH820899	BACBP15-A
CC858148	NDL.129K9
AQ990998	RPCT-23-3
CK818278	ic23f06.y
CT551967	OB.BE003
AU217527	AU217527
BT702183	BT702183
AG331381	Mus muscu
CF264610	AGENCOURT
BZ847658	CH240.250
BT815333	BT815333
CE770220	LGRT-888-
BH487731	BOHUB9TF
AZ210199	SP.0153.A
CK026066	AGENCOURT
CC816891	100002X06
CA474062	AGENCOURT
CC133823	NDL.28P16
AZ136157	SP.0169.A
CL022024	CH216-13B
BX285738	Arabidops
AL946816	Arabidops
A1626105	ar0801.x
CR401435	Arabidops
BG211404	RSR30966
D72160	CELRK0906ZR
BB233201	BB233201
CN924215	00041AEL
CN453602	E03-CK-L2
AQ194703	RPCT11-59
BX221357	Danio rer
AQ410152	HS.5083.B
AU061419	AU061419
BM599153	17006590
AJ672916	AJ672916
CB909797	USDA-PP.1
CA865978	136d12.x
BX285737	Arabidops
CD450045	USDA-PP.1
AQ829583	HS.4832.A
BJ435871	BJ435871
A1489028	EST247367
AQ661945	HS.2093.B
BF003557	EST432055
BM359958	GA.EB002
AQ593197	HS.5465.A
CD842212	RF02.1261
CE128067	LGRT-988-
CD195244	MS1-0088T
AQ155496	HS.3055.B
AQ245575	HS.2037.B
BH527994	BH527994
BZ347068	hwo8a01.B
BZ351646	hwo8a01.B
AQ130690	HS.3009.A
CA841028	1532c07.x
CK523454	reweao.00
BX237955	Danio rer
AZ925835	476.dic03
CNS00VM9	Arabidops
AZ927695	476.dic28
CO276626	EX121657.
B61330	T22C7TR.TAM
AQ060625	HS.5404.A
AQ060625	HS.5447.A
AQ565565	HS.2110.A
BJ482831	BJ482831
AU262981	AU262981

C 98 16 64.0 562 8 A0172127 HS 3170 B
C 98 16 64.0 563 1 A038197 AU038197
100 16 64.0 565 4 B156511 B156511 RA43196.5

ALIGNMENTS

RESULT 1
LOCUS CC141515 877 bp DNA linear GSS 16-APR-2003
DEFINITION NDL.32K2.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
ACCESSION NDL.32K2, genomic survey sequence.
VERSION CC141515
KEYWORDS CC141515.1 GI:30010570
SOURCE GSS.
ORGANISM Aedes aegypti (yellow fever mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes;
Stegomyia.

REFERENCE 1 (bases 1 to 877)
AUTHORS Loftus, B., Shetty, J., Knudson, D. and Severson, D.
TITLE BAC end sequencing of Aedes aegypti
JOURNAL Unpublished (2003)
COMMENT Other GSSs: NDL.32K2.T7
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: SP6
Class: BAC ends.

FEATURES

source
1..877
/location/Qualifiers
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="NDL.32K2"
/clone_lib="Notre Dame Liverpool"
/note="Vector: pECBAC1; Site_1: Hind III. The library was prepared from whole body tissue of newly hatched 1d larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 76.0%; Score 19; DB 8; Length 877;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 GAATAATGAAATATATCT 23
DB 575 GAATATTTGAATATATCT 593

RESULT 2
LOCUS CG965204 888 bp DNA linear GSS 15-DEC-2003
DEFINITION MBEMD88TRF mth2 Medicago truncatula genomic clone 84P8, genomic survey sequence.
ACCESSION CG965204
VERSION CG965204.1 GI:39890624
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.

REFERENCE 1 (bases 1 to 888)
AUTHORS Town, C.D., Shetty, J., Koo, H. and Feldblum, T.F.
TITLE Sequencing of BAC ends from Medicago truncatula
JOURNAL Unpublished (2003)
COMMENT Other GSSs: MBEMD88TR
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
Seq primer: TGTATATACGACGGCCAGT
Class: BAC ends.

FEATURES

source
1..888
/location/Qualifiers
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="genotype A17"
/db_xref="taxon:3850"
/clone="84P8"
/clone_lib="mth2"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII; Cook, D.R. and Kim, D.J. unpublished"

ORIGIN

Query Match 76.0%; Score 19; DB 9; Length 888;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 GAATAATGAAATATATCT 23
DB 129 GAATATTTGAATATATCT 147

RESULT 3
LOCUS CR286805 668 bp mRNA linear EST 27-FEB-2004
DEFINITION CR286805 Oryza sativa library (Han B) Oryza sativa cDNA clone
ACCESSION P720G08p3, mRNA sequence.
VERSION CR286805
KEYWORDS CR286805.1 GI:44673371
SOURCE EST.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Emhaltoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 668)
Han, B., Feng, Q., Huang, Y.C., Yang, K., Li, Y., Guan, J.P., Zhu, J.J., Zhao, Q., Hu, X., Liu, Y.L., Ma, J., Yu, Z., Chen, L., Fan, D.L., Wang, Q.J., Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Zhang, L., Fan, L.F., Chen, W., Wu, S.A. and Xue, Y.B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhanengr.ac.cn
Clone requests: bhanengr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.
location/Qualifiers

FEATURES

source
1..668
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="P720G08p3"
/clone_lib="Oryza sativa library (Han B)"

ORIGIN

Query Match 72.0%; Score 18; DB 7; Length 668;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:28:49 : Search time 43.018 Seconds
(without alignments)
3138.794 Million cell updates/sec

Title: US-10-088-666-73

Perfect score: 25

Sequence: 1 tcgagataatcgataatcattag 25

Scoring table: Oligo NUC, Gapop 60.0, Gapext 60.0

Searched: 3625171 seqs, 2700493622 residues

Word size: 10

Total number of hits satisfying chosen parameters: 80234

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 100 summaries

Database: Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	18	72.0	1190	9	US-09-070-927A-714
C 2	18	72.0	1891	17	US-10-437-563-45884
C 3	18	72.0	5962	15	US-10-311-455-1259
C 4	17	60.0	317	17	US-10-767-795-962
C 5	17	60.0	672	16	US-10-282-122A-31065
C 6	17	60.0	1283	17	US-10-767-795-961
C 7	16	64.0	334	16	US-10-242-535A-13072
C 8	16	64.0	334	16	US-10-085-783A-13072
C 9	16	64.0	339	16	US-10-242-535A-13028
C 10	16	64.0	339	16	US-10-085-783A-13028
C 11	16	64.0	442	16	US-10-424-599-140935
C 12	16	64.0	455	9	US-09-864-761-5127

C 13	16	64.0	817	13	US-10-027-633-155971	Sequence 155971, A
C 14	16	64.0	817	13	US-10-027-633-155971	Sequence 155971, A
C 15	16	64.0	949	13	US-10-027-633-12128	Sequence 12128, A
C 16	16	64.0	949	13	US-10-027-633-12128	Sequence 12128, A
C 17	16	64.0	950	13	US-10-027-633-254585	Sequence 254585, A
C 18	16	64.0	950	13	US-10-027-633-254585	Sequence 254585, A
C 19	16	64.0	2062	13	US-10-027-633-257250	Sequence 257250, A
C 20	16	64.0	2062	13	US-10-027-633-257250	Sequence 257250, A
C 21	16	64.0	2136	16	US-10-282-122A-21490	Sequence 21490, A
C 22	16	64.0	2415	16	US-10-282-122A-21490	Sequence 21490, A
C 23	16	64.0	6527	15	US-10-356-631-3	Sequence 31273, A
C 24	16	64.0	8979	16	US-10-311-455-758	Sequence 3, Appl.1
C 25	16	64.0	8979	16	US-10-221-613-138	Sequence 758, App
C 26	16	64.0	38646	15	US-10-085-117-193	Sequence 138, App
C 27	16	64.0	78695	18	US-10-085-117-193	Sequence 193, App
C 28	16	64.0	322101	15	US-10-060-902-1	Sequence 257, App
C 29	16	64.0	322101	15	US-10-060-902-1	Sequence 1, Appl.1
C 30	16	64.0	3673778	15	US-10-354-243-1	Sequence 1, Appl.1
C 31	15	60.0	220	16	US-10-312-841-1	Sequence 9107, Ap
C 32	15	60.0	220	16	US-10-242-535A-9107	Sequence 9107, Ap
C 33	15	60.0	220	16	US-10-085-783A-9107	Sequence 139611, A
C 34	15	60.0	273	18	US-10-425-115-132611	Sequence 31354, A
C 35	15	60.0	313	16	US-10-242-535A-13354	Sequence 31354, A
C 36	15	60.0	313	16	US-10-085-783A-13354	Sequence 1054, Ap
C 37	15	60.0	326	9	US-09-983-965-1054	Sequence 28337, A
C 38	15	60.0	449	17	US-10-767-701-28337	Sequence 40912, A
C 39	15	60.0	466	13	US-10-027-633-40912	Sequence 40912, A
C 40	15	60.0	466	13	US-10-027-633-40912	Sequence 96239, A
C 41	15	60.0	520	13	US-10-027-633-96239	Sequence 307284, A
C 42	15	60.0	520	13	US-10-027-633-96239	Sequence 96239, A
C 43	15	60.0	520	13	US-10-027-633-307284	Sequence 307284, A
C 44	15	60.0	529	9	US-09-860-670-28	Sequence 28, Appl.1
C 45	15	60.0	529	9	US-10-027-633-4343	Sequence 43, Ap
C 46	15	60.0	535	15	US-10-029-366-4343	Sequence 65, Appl.1
C 47	15	60.0	572	11	US-10-315-664-43	Sequence 188472, A
C 48	15	60.0	572	11	US-10-315-664-43	Sequence 188472, A
C 49	15	60.0	579	13	US-10-027-633-188472	Sequence 41988, A
C 50	15	60.0	579	13	US-10-027-633-41988	Sequence 41988, A
C 51	15	60.0	583	13	US-10-027-633-41988	Sequence 62113, A
C 52	15	60.0	583	13	US-10-027-633-41988	Sequence 41988, A
C 53	15	60.0	583	13	US-10-027-633-41988	Sequence 62113, A
C 54	15	60.0	583	13	US-10-027-633-41988	Sequence 62113, A
C 55	15	60.0	634	15	US-10-027-633-244713	Sequence 244713, A
C 56	15	60.0	634	15	US-10-027-633-244713	Sequence 70391, A
C 57	15	60.0	633	13	US-10-027-633-70391	Sequence 71677, A
C 58	15	60.0	633	13	US-10-027-633-71677	Sequence 294286, A
C 59	15	60.0	643	13	US-10-027-633-71677	Sequence 110991, A
C 60	15	60.0	643	13	US-10-027-633-71677	Sequence 110991, A
C 61	15	60.0	643	13	US-10-027-633-71677	Sequence 3211, Ap
C 62	15	60.0	643	13	US-10-027-633-71677	Sequence 1789, Ap
C 63	15	60.0	643	13	US-10-027-633-71677	Sequence 138846, A
C 64	15	60.0	643	13	US-10-027-633-71677	Sequence 138846, A
C 65	15	60.0	644	13	US-10-027-633-294286	Sequence 67, Appl.1
C 66	15	60.0	644	13	US-10-027-633-294286	Sequence 159755, A
C 67	15	60.0	644	13	US-10-027-633-294286	Sequence 159755, A
C 68	15	60.0	644	13	US-10-027-633-294286	Sequence 159755, A
C 69	15	60.0	786	15	US-10-106-698-1789	Sequence 7194, Ap
C 70	15	60.0	806	15	US-10-027-633-1789	Sequence 7194, Ap
C 71	15	60.0	806	15	US-10-027-633-1789	Sequence 7194, Ap
C 72	15	60.0	854	11	US-09-903-190-67	Sequence 159756, A
C 73	15	60.0	854	11	US-09-903-190-67	Sequence 159756, A
C 74	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 75	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 76	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 77	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 78	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 79	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 80	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 81	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 82	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 83	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 84	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A
C 85	15	60.0	856	13	US-10-027-633-159756	Sequence 159756, A

```

c 86 15 60.0 1072 16 US-10-424-599-114177 Sequence 114177,
c 87 15 60.0 1152 18 US-10-425-115-93212 Sequence 93212, A
c 88 15 60.0 1166 9 US-09-860-670-231 Sequence 231, App
c 89 15 60.0 1166 15 US-10-227-646-231 Sequence 231, App
c 90 15 60.0 2000 9 US-09-938-842A-3288 Sequence 3288, App
c 91 15 60.0 2000 11 US-09-938-842A-3288 Sequence 3288, App
c 92 15 60.0 2136 9 US-09-860-670-227 Sequence 227, App
c 93 15 60.0 2136 15 US-10-227-646-227 Sequence 227, App
c 94 15 60.0 2275 16 US-10-425-114-6036 Sequence 6036, App
c 95 15 60.0 2311 17 US-10-437-963-6048 Sequence 6048, App
c 96 15 60.0 2581 16 US-10-429-949-187 Sequence 187, App
c 97 15 60.0 3001 15 US-10-147-603-187 Sequence 187, App
c 98 15 60.0 3005 18 US-10-425-115-52901 Sequence 52901, A
c 99 15 60.0 3519 16 US-10-282-122A-21681 Sequence 21681, A
c 100 15 60.0 3927 15 US-10-252-157-260 Sequence 260, App

```

ALIGNMENTS

```

RESULT 1
US-09-070-927A-714/C
; Sequence 714, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunach
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070, 927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 714:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1190 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 714:
US-09-070-927A-714

```

```

Query Match 72.0%; Score 18; DB 9; Length 1190;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 3 GAGATTAATTGATATA 20
Db 396 GAGATTAATTGATATA 379

```

```

RESULT 2
US-10-437-963-45894/C
; Sequence 45894, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 45894
; LENGTH: 1891
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48814C.1
US-10-437-963-45894

```

```

Query Match 72.0%; Score 18; DB 17; Length 1891;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 6 AATAATTGATATAATCT 23
Db 1793 AATAATTGATATAATCT 1776

```

```

RESULT 3
US-10-311-455-1259
; Sequence 1259, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1259
; LENGTH: 5962
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 5264
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1259

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:18:54 ; Search time 10.6607 Seconds

(without alignments)
1666,850 Million cell updates/sec

Title: US-10-088-666-73

Perfect score: 25

Sequence: 1 tcgagataatgataatcattag 25

Scoring table: Oligo_NUC

Gapop 60.0 ; Gapext 60.0

Searched: 824507 seqs, 355394441 residues

Word size: 10

Total number of hits satisfying chosen parameters: 12495

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Issued Patents, NA:*

1: /cgn2_6/ptodata/1/ina/5A_CONB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_CONB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_CONB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_CONB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUTS_CONB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	64.0	183	4	US-09-270-767-27323
2	16	64.0	442	4	US-09-270-767-11703
3	16	64.0	1275	4	US-09-248-796A-3349
4	16	64.0	6527	4	US-09-492-308A-3
5	16	60.0	285	4	US-09-248-796A-11887
6	15	60.0	492	4	US-09-621-976-18621
7	15	60.0	496	4	US-09-621-976-1493
8	15	60.0	572	4	US-09-529-360B-65
9	15	60.0	854	4	US-09-247-155-67
10	15	60.0	3001	4	US-09-539-333D-187
11	15	60.0	3552	4	US-09-107-532A-1026
12	15	60.0	5064	4	US-08-936-107A-8
13	15	60.0	5923	3	US-09-064-922-3
14	15	60.0	112132	4	US-09-741-150-3
15	15	60.0	112132	4	US-10-160-187-3
16	14	56.0	183	4	US-09-248-796A-9851
17	14	56.0	183	4	US-09-248-796A-10090
18	14	56.0	273	4	US-09-543-681A-211
19	14	56.0	394	4	US-08-956-171E-909
20	14	56.0	394	4	US-08-781-986A-909
21	14	56.0	417	4	US-09-543-681A-643
22	14	56.0	462	4	US-09-513-999C-22848
23	14	56.0	570	4	US-09-248-796A-14078
24	14	56.0	623	4	US-09-270-767-15116
25	14	56.0	657	4	US-09-248-796A-11566
26	14	56.0	744	4	US-09-248-796A-2514
27	14	56.0	759	4	US-09-710-279-1031

C 28	14	56.0	969	4	US-09-248-796A-1329	Sequence 1329, Ap
C 29	14	56.0	1296	4	US-09-710-279-1473	Sequence 1473, Ap
C 30	14	56.0	1353	3	US-09-134-001C-1974	Sequence 1974, Ap
C 31	14	56.0	1443	3	US-09-221-017B-685	Sequence 685, Ap
C 32	14	56.0	1452	4	US-09-107-532A-2242	Sequence 2242, Ap
C 33	14	56.0	1506	3	US-09-134-001C-1278	Sequence 1278, Ap
C 34	14	56.0	1743	3	US-09-134-001C-2774	Sequence 2774, Ap
C 35	14	56.0	1816	4	US-08-956-171E-351	Sequence 351, App
C 36	14	56.0	1816	4	US-08-781-986A-331	Sequence 331, App
C 37	14	56.0	1972	1	US-08-463-048-1	Sequence 1, Appli
C 38	14	56.0	1972	1	US-08-463-229-1	Sequence 1, Appli
C 39	14	56.0	1972	2	US-08-302-891-1	Sequence 1, Appli
C 40	14	56.0	2007	4	US-09-248-796A-4799	Sequence 4799, Ap
C 41	14	56.0	2150	4	US-09-248-796A-3599	Sequence 3599, Ap
C 42	14	56.0	2220	4	US-09-543-681A-2265	Sequence 2265, Ap
C 43	14	56.0	2250	4	US-09-328-352-3462	Sequence 3462, Ap
C 44	14	56.0	2301	4	US-09-710-279-419	Sequence 419, App
C 45	14	56.0	2441	1	US-08-920-812-12	Sequence 12, Appl
C 46	14	56.0	2441	1	US-08-920-827-12	Sequence 12, Appl
C 47	14	56.0	2441	1	US-08-921-177-12	Sequence 12, Appl
C 48	14	56.0	2441	1	US-08-920-828-12	Sequence 12, Appl
C 49	14	56.0	2634	2	US-09-110-279-4189	Sequence 4189, Ap
C 50	14	56.0	2634	4	US-09-710-279-4069	Sequence 4069, Ap
C 51	14	56.0	2744	4	US-09-489-847-111	Sequence 111, App
C 52	14	56.0	2871	4	US-09-489-847-22	Sequence 22, Appl
C 53	14	56.0	2876	4	US-09-489-847-22	Sequence 22, Appl
C 54	14	56.0	3000	4	US-09-710-279-4422	Sequence 4422, Ap
C 55	14	56.0	3017	4	US-09-710-279-3882	Sequence 3882, Ap
C 56	14	56.0	3170	4	US-09-710-279-3867	Sequence 3867, Ap
C 57	14	56.0	3229	4	US-09-710-279-3566	Sequence 3566, Ap
C 58	14	56.0	3233	4	US-09-710-279-4058	Sequence 4058, Ap
C 59	14	56.0	3321	4	US-09-710-279-4241	Sequence 4241, Ap
C 60	14	56.0	3438	4	US-09-710-279-3430	Sequence 3430, Ap
C 61	14	56.0	3602	4	US-09-710-279-3799	Sequence 3799, Ap
C 62	14	56.0	3825	4	US-09-710-279-3641	Sequence 3641, Ap
C 63	14	56.0	4094	4	US-09-710-279-3646	Sequence 3646, Ap
C 64	14	56.0	4185	4	US-10-204-708-12	Sequence 12, Appl
C 65	14	56.0	6317	1	US-08-468-036-4	Sequence 4, Appli
C 66	14	56.0	6854	2	US-08-376-843-4	Sequence 759, App
C 67	14	56.0	6984	2	US-09-134-001C-627	Sequence 627, App
C 68	14	56.0	6984	3	US-09-710-279-759	Sequence 759, App
C 69	14	56.0	7215	3	US-09-134-001C-627	Sequence 627, App
C 70	14	56.0	9072	4	US-08-956-171E-45	Sequence 45, Appl
C 71	14	56.0	9072	4	US-08-781-986A-45	Sequence 45, Appl
C 72	14	56.0	99500	3	US-09-798-096-10	Sequence 10, Appl
C 73	14	56.0	128779	4	US-09-497-855A-38	Sequence 38, Appl
C 74	14	56.0	640681	4	US-09-790-988-1	Sequence 1, Appli
C 75	14	56.0	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 76	14	56.0	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 77	14	56.0	1830121	4	US-10-329-960-1	Sequence 1, Appli
C 78	13	52.0	111	4	US-09-722-219-4	Sequence 4, Appli
C 79	13	52.0	129	4	US-09-722-219-11	Sequence 11, Appl
C 80	13	52.0	165	4	US-09-722-219-2	Sequence 2, Appli
C 81	13	52.0	186	4	US-09-513-999C-2761	Sequence 2761, Ap
C 82	13	52.0	199	4	US-09-513-999C-28284	Sequence 28284, A
C 83	13	52.0	209	4	US-09-621-976-8675	Sequence 8675, Ap
C 84	13	52.0	222	4	US-09-621-976-8105	Sequence 8105, Ap
C 85	13	52.0	244	4	US-09-621-976-8105	Sequence 8105, Ap
C 86	13	52.0	240	4	US-09-248-796A-8796	Sequence 8796, Ap
C 87	13	52.0	246	4	US-09-543-681A-2795	Sequence 2795, Ap
C 88	13	52.0	246	4	US-09-134-000C-183	Sequence 1383, Ap
C 89	13	52.0	252	4	US-09-248-796A-13268	Sequence 13268, A
C 90	13	52.0	280	4	US-09-513-999C-3766	Sequence 3766, Ap
C 91	13	52.0	338	4	US-09-621-976-9471	Sequence 9471, Ap
C 92	13	52.0	338	4	US-09-621-976-803	Sequence 803, Ap
C 93	13	52.0	345	4	US-09-621-976-8620	Sequence 8620, Ap
C 94	13	52.0	345	4	US-09-248-796A-13197	Sequence 13197, A
C 95	13	52.0	348	4	US-09-621-976-884	Sequence 884, App
C 96	13	52.0	350	4	US-09-621-976-9498	Sequence 9498, Ap
C 97	13	52.0	352	4	US-09-621-976-19218	Sequence 19218, A
C 98	13	52.0	354	4	US-08-956-171E-1667	Sequence 1467, Ap
C 99	13	52.0	354	4	US-08-781-986A-167	Sequence 1467, Ap
C 100	13	52.0	359	4	US-09-621-976-8657	Sequence 8657, Ap

ALIGNMENTS

RESULT 1

US-09-270-767-27323
 ; Sequence 27323, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 27323
 ; LENGTH: 183
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 ; US-09-270-767-27323

Query Match 64.0%; Score 16; DB 4; Length 183;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAATATATCTA 24
 |||||
 Db 85 AATTGAATATATCTA 100

RESULT 2

US-09-270-767-11703
 ; Sequence 11703, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 11703
 ; LENGTH: 442
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 ; US-09-270-767-11703

Query Match 64.0%; Score 16; DB 4; Length 442;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAATATATCTA 24
 |||||
 Db 344 AATTGAATATATCTA 359

US-09-248-796A-3349/c

US-09-248-796A-3349/c
 ; Sequence 3349, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstein et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409

US-09-248-796A-3349
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO: 3349
 ; LENGTH: 1275
 ; TYPE: DNA
 ; ORGANISM: Candida albicans

Query Match 64.0%; Score 16; DB 4; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 3.8;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAATATATCTA 24
 |||||
 Db 203 AATTGAATATATCTA 188

RESULT 4

US-09-492-308A-3
 ; Sequence 3, Application US/09492308A
 ; Patent No. 6573430
 ; GENERAL INFORMATION:
 ; APPLICANT: Bradley, DJ
 ; Carpenter, R
 ; Coen, ES

TITLE OF INVENTION: Flowering genes
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Nixon & Vanderhye PC
 STREET: 8th Floor, 1100 No. 6573430th Glebe Road
 CITY: Arlington
 STATE: Virginia
 COUNTRY: United States of America
 ZIP: 22201-4714

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/492,308A
 FILING DATE: 27-Jan-2000
 CLASSIFICATION: 435 (Preliminary)

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/043,029
 FILING DATE: 13-MAR-1998
 APPLICATION NUMBER: PCT/GB96/02276
 FILING DATE: 13-SEP-1996
 APPLICATION NUMBER: GB 9518731.6
 FILING DATE: 13-SEP-1995

ATTORNEY/AGENT INFORMATION:
 NAME: Mary J. Wilson
 REGISTRATION NUMBER: 32,955
 REFERENCE/DOCKET NUMBER: 620-93
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6527 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single

TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Antirrhinum majus
 IMMEDIATE SOURCE:
 CLONE: CEN GENOMIC
 POSITION IN GENOME:

UNITS: bp
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-09-492-308A-3

c 93 17 68.0 196510 2 AC141117
 94 17 68.0 197079 2 AC147385
 95 17 68.0 199326 2 AC120976
 c 96 17 68.0 201198 2 AC114616
 97 17 68.0 202587 2 AC119862
 c 98 17 68.0 203266 2 AC097006
 99 17 68.0 203737 10 AC116461
 100 17 68.0 208225 10 AC111022

ALIGNMENTS

RESULT 1
 AX105801 25 bp DNA linear PAT 30-APR-2001
 LOCUS AX105801
 DEFINITION Sequence 73 from Patent WO0123605.
 ACCESSION AX105801
 VERSION AX105801.1 GI:13921814
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 Pandke, M., Gasch, A. and Berghof, K.
 Method and nucleic acids for determining the presence of
 micro-organisms specific to the brewing process
 Patent: WO 0123605-A 73 05-APR-2001;
 Biorecon Diagnostics GmbH (DE)
 location/Qualifiers

FEATURES
 source
 1..25
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="spezifische Sequenz fur Lactobacillus brevis"

ORIGIN
 Query Match 100.0%; Score 25; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATATCTAG 25
 Db 1 TCGAGATATATGATATATATCTAG 25

RESULT 2
 AX105729 267 bp DNA linear PAT 30-APR-2001
 LOCUS AX105729
 DEFINITION Sequence 1 from Patent WO0123605.
 ACCESSION AX105729
 VERSION AX105729.1 GI:13921742
 KEYWORDS
 SOURCE
 ORGANISM

Lactobacillus brevis
 Lactobacillus brevis
 Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 Lactobacillus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1..267
 /organism="Lactobacillus brevis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1580"

ORIGIN
 Query Match 100.0%; Score 25; DB 6; Length 267;
 Best Local Similarity 100.0%; Pred. No. 0.0092;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATATCTAG 25
 Db 125 TCGAGATATATGATATATATCTAG 149

RESULT 3
 AC079132/c 64701 bp DNA linear HTG 23-FEB-2001
 LOCUS AC079132
 DEFINITION Homo sapiens chromosome 8 clone RP11-214U7 map 8, LOW-PASS SEQUENCE
 SAMPLING.
 AC079132
 AC079132.2 GI:13112249
 HTG: HTGS PHASEO.
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 8, clone RP11-214U7
 Unpublished
 2 (bases 1 to 64701)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Barta, N., Bastien, V., Beda, F., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Burkett, G., Campoliano, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 Dearliano, K., Dewar, K., Diaz, J.S., Dodg, S., Ferreira, P.,
 Fitzhugh, W., Gage, D., Galagan, J., Gardys, S., Ginde, S., Goyette, M.,
 Graham, L., Grand-Pierre, N., Hagos, B., Hesford, A., Horton, L.,
 Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Labocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lien, C., Liu, G.,
 MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McKernan, K.,
 McPheeters, R., Meldrum, D., Menus, L., Mohova, T., Mlenga, V.,
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
 O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
 Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
 Rogov, P., Rottman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
 Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, A., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Tirelli, A., Travers, M., Triggilo, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zimmer, A. and Zody, M.

COMMENT
 JOURNAL
 TITLE
 JOURNAL
 COMMENT
 Submitted (19-AUG-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 23, 2001 this sequence version replaced gi:19857544.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WISR
 Web site: http://www-seg.wi.mit.edu
 Contact: sequence submissions@genome.wi.mit.edu
 Project Information
 Center project name: I10419
 Center clone name: 214_U_7

NOTE: This record contains 81 individual
 sequencing reads that have not been assembled into
 contigs. Runs of N are used to separate the reads
 and the order in which they appear is completely
 arbitrary. Low-pass sequence sampling is useful for
 identifying clones that may be gene-rich and allows
 overlap relationships among clones to be deduced.
 However, it should not be assumed that this clone
 will be sequenced to completion. In the event that
 the record is updated, the accession number will
 be preserved.

1 661: contig of 661 bp in length
 662 761: gap of 100 bp
 762 1473: contig of 712 bp in length

Tue Nov 16 16:09:52 2004

us-10-088-666-21.olig10.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:16:19 ; Search time 121.396 Seconds
(without alignments)
6303.600 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21

Sequence: 1 ccaagtcacacacgtagtgc 21

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 1821986598 residues

Word size: 10

Total number of hits satisfying chosen parameters: 248193

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_nuc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	85.7	779	8	BH542735 BOGXU95TR
2	17	81.0	727	8	BH225601 603401079
3	17	81.0	796	9	CG361559 CG1DT94TH
4	17	81.0	797	7	CK096311 UB10CP810
5	16	76.2	228	9	CR349126 Medicago
6	16	76.2	544	9	CG467736 ZMMEB025
7	16	76.2	588	8	BZ294416 CG0884.f1
8	16	76.2	648	9	CG307774 CGXR17TH
9	16	76.2	674	9	CG307783 CGXR17TH
10	16	76.2	682	5	BM213486 BM213486
11	16	76.2	702	8	BH179655 BH179655
12	16	76.2	702	8	CGN60710T T7 end of
13	16	76.2	709	8	BH926500 od142f09.
14	16	76.2	789	9	CG883248 ZMMEB049
15	16	76.2	797	9	CG390193 ZMMEB051
16	16	76.2	830	8	BZ076210 1KFE1D10.
17	16	76.2	894	9	CG044759 PUF8X79TD
18	16	76.2	940	9	CG044756 PUF8X79TB
19	16	76.2	959	9	CG609999 CGULH26TH
20	16	76.2	985	9	CG610004 CGULH26TV
21	16	76.2	1022	8	CG370241 PUBHY44TB
22	16	76.2	1063	9	CL202290 ZMMEB056
23	15	71.4	244	7	CN723309 tae63c01.
24	15	71.4	260	9	CL223183 ZMMEB050

C	25	15	71.4	291	6	CAB54679	CAB54679
C	26	15	71.4	326	7	CF542671	CF542671
C	27	15	71.4	342	8	CC400403	CC400403
C	28	15	71.4	377	1	AV203663	AV203663
C	29	15	71.4	410	6	CB102538	CB102538
C	30	15	71.4	420	9	CC719572	CC719572
C	31	15	71.4	443	7	CK581397	CK581397
C	32	15	71.4	451	7	CK583556	CK583556
C	33	15	71.4	480	1	AU168090	AU168090
C	34	15	71.4	480	5	BP047360	BP047360
C	35	15	71.4	501	5	BP072186	BP072186
C	36	15	71.4	505	1	AV767679	AV767679
C	37	15	71.4	517	5	BP040971	BP040971
C	38	15	71.4	521	7	CK583484	CK583484
C	39	15	71.4	522	7	CK583669	CK583669
C	40	15	71.4	532	7	CK583445	CK583445
C	41	15	71.4	535	4	BM275391	BM275391
C	42	15	71.4	535	4	BM347245	BM347245
C	43	15	71.4	537	7	CN046315	CN046315
C	44	15	71.4	538	7	CF542672	CF542672
C	45	15	71.4	548	2	AW734365	AW734365
C	46	15	71.4	557	7	CK583684	CK583684
C	47	15	71.4	559	6	CAB56722	CAB56722
C	48	15	71.4	566	6	CB102087	CB102087
C	49	15	71.4	568	4	BJ122833	BJ122833
C	50	15	71.4	568	4	BJ122973	BJ122973
C	51	15	71.4	571	1	AU200886	AU200886
C	52	15	71.4	574	8	CC345942	CC345942
C	53	15	71.4	578	7	CK583435	CK583435
C	54	15	71.4	579	6	CAB55682	CAB55682
C	55	15	71.4	580	4	BJ108163	BJ108163
C	56	15	71.4	585	6	CB102412	CB102412
C	57	15	71.4	587	6	CAB54964	CAB54964
C	58	15	71.4	606	6	CB102156	CB102156
C	59	15	71.4	617	4	BG840260	MEST11-F0
C	60	15	71.4	619	5	BM351601	BM351601
C	61	15	71.4	634	4	CK581081	CK581081
C	62	15	71.4	645	4	EM584859	EM584859
C	63	15	71.4	645	5	BM336644	BM336644
C	64	15	71.4	657	9	AG068562	Pan trogl
C	65	15	71.4	673	7	CK581122	CK581122
C	66	15	71.4	678	1	AV841373	AV841373
C	67	15	71.4	681	5	BM075694	BM075694
C	68	15	71.4	684	9	CG391652	ZMMEB057
C	69	15	71.4	712	8	EZ685709	PUBDF15TD
C	70	15	71.4	721	6	CB102546	MMV_SQ010
C	71	15	71.4	743	9	AG446717	Mus muscu
C	72	15	71.4	745	9	CC786909	ZMMEB015
C	73	15	71.4	753	8	CC400402	PURTH08TB
C	74	15	71.4	757	9	CL447328	ZMMEB046
C	75	15	71.4	770	5	BM350773	BM350773
C	76	15	71.4	774	5	BM401388	BM401388
C	77	15	71.4	790	9	CC676527	OGFT09TH
C	78	15	71.4	807	5	BO144662	NFO76R03D
C	79	15	71.4	825	8	BZ639321	CCCTB3TM
C	80	15	71.4	829	7	CK453913	911160 MA
C	81	15	71.4	840	7	CC035870	EST814254
C	82	15	71.4	852	7	CF812632	EST696014
C	83	15	71.4	871	8	CC436559	PURLH79TD
C	84	15	71.4	882	7	CC016802	EST785184
C	85	15	71.4	892	7	CO014788	EST785170
C	86	15	71.4	903	9	CG325319	OC2CD17TH
C	87	15	71.4	913	9	CF814429	EST651811
C	88	15	71.4	915	7	CO017225	EST787607
C	89	15	71.4	919	5	BU421326	603954553
C	90	15	71.4	958	7	CF818080	EST695462
C	91	15	71.4	1008	7	CC0032335	EST810719
C	92	15	71.4	1054	9	CNS06YED	T3 end of
C	93	14	66.7	182	4	AG156582	8ab11911.
C	94	14	66.7	200	2	AW788050	CO0148-R
C	95	14	66.7	251	4	AA259220	MA1MR043.
C	96	14	66.7	251	1	BI863831	kx48n09.y
C	97	14	66.7	256	1	AI093870	ga30e08.s

C 98 14 66.7 276 9 C5590496 tigr-gss-
 98 14 66.7 288 1 AA791342 SMTBADA00
 C 100 14 66.7 328 7 C0750077 KBFAED04P

ALIGNMENTS

RESULT 1
 BH542735 779 bp DNA linear GSS 14-DEC-2001
 LOCUS BOGXU95TR BOGX Brassica oleracea genomic clone BOGXU95, genomic
 DEFINITION survey sequence.

ACCESSION BH542735
 VERSION BH542735.1 GI:17794516
 KEYWORDS

SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 779)
 Town, C.D., Van Aken, S., Uteerback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea

AUTHORS Town, C.D., Van Aken, S., Uteerback, T., Koo, H. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished (2001)
 COMMENT Other GSSs: BOGXU95TR
 Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES

source
 1..779
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOGXU95"
 /clone_1ib="BOGX"
 /note="Vector: pROSL; Site_1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pROSL using BstXI linkers"

ORIGIN

Query Match 85.7%; Score 18; DB 8; Length 779;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CAAGTCACACGCTAGT 19
 Db 356 CAAGTCACACGCTAGT 373

RESULT 2
 BU225601/c 727 bp mRNA linear EST 26-NOV-2002
 LOCUS 603401079F1 CSEQCHN23 Gallus gallus cDNA clone CSEST955nl9 5', mRNA
 DEFINITION sequence.

ACCESSION BU225601
 VERSION BU225601.1 GI:25461670
 KEYWORDS EST
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 727)
 Boardman, P.E., Sanz-Ruiz, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.O.

AUTHORS Boardman, P.E., Sanz-Ruiz, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.O.
 TITLE A Comprehensive Collection of Chicken CDVs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 2233534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)

PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source
 1..727
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CSEST955nl9"
 /dev_stage="22"
 /lab_host="DH10B"
 /clone_1ib="CSEQCHN23"
 /note="Organ: head; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this, first strand reaction, double-stranded cDNA
 was bluntended, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldi et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Query Match 81.0%; Score 17; DB 5; Length 727;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGTCACACGCTAG 17
 Db 91 CCAAGTCACACGCTAG 75

RESULT 3
 CG361559

LOCUS CG361559 796 bp DNA linear GSS 26-AUG-2003
 DEFINITION CG1DTP94TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMEWA0739020,
 genomic survey sequence.

ACCESSION CG361559
 VERSION CG361559.1 GI:34278826
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE 1 (bases 1 to 796)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

White, C.A., Quackenbush, J., Van Aken, S., Uteerback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, V.A., Kohlring, T.,
 Citek, R.W., Numborg, A., Robbins, D. and Lakey, N.

CONSORTIUM for Maize Genomics
 Unpublished (2002)
 Other_GSSs: CG1DTP94TV

COMMENT Contact: Cathy Whitelaw
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:59:59 ; Search time 48.4234 Seconds
(without alignments)
2710.166 Million cell updates/sec

Title: US-10-088-666-73

Perfect score: 25

Sequence: 1 tcgagaatcatgatatatctag 25

Scoring table: **OLIGO-NUC**
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size: 10

Total number of hits satisfying chosen parameters: 72573

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	5	AA61638 Lactobaci
2	25	100.0	267	5	AA61638 Lactobaci
3	18	72.0	1190	2	AA61638 Lactobaci
4	18	72.0	1190	6	AA61638 Lactobaci
5	18	72.0	5962	6	AA61638 Lactobaci
6	17	68.0	672	8	AA61638 Lactobaci
7	17	68.0	177531	8	AA61638 Lactobaci
8	17	68.0	177531	10	AA61638 Lactobaci
9	17	68.0	177531	10	AA61638 Lactobaci
10	17	68.0	177531	10	AA61638 Lactobaci
11	17	68.0	177531	10	AA61638 Lactobaci
12	17	68.0	177531	10	AA61638 Lactobaci
13	17	68.0	177531	10	AA61638 Lactobaci
14	17	68.0	177531	10	AA61638 Lactobaci
15	17	68.0	177531	10	AA61638 Lactobaci
16	17	68.0	177531	10	AA61638 Lactobaci
17	17	68.0	177531	10	AA61638 Lactobaci
18	17	68.0	177531	10	AA61638 Lactobaci
19	17	68.0	177531	10	AA61638 Lactobaci
20	17	68.0	177531	10	AA61638 Lactobaci
21	17	68.0	177531	10	AA61638 Lactobaci

22	16	64.0	3309	4	ABL21530
23	16	64.0	6527	2	AA61638
24	16	64.0	8979	6	ABL32765
25	16	64.0	8979	6	ABL31271
26	16	64.0	8979	6	ABL70232
27	16	64.0	8979	6	ABL61179
28	16	64.0	25426	8	ADA1645
29	16	64.0	25426	8	ACC50940
30	16	64.0	25426	10	ADD38155
31	16	64.0	25426	10	ADA57777
32	16	64.0	110000	10	ACR67367
33	16	64.0	110000	10	ACR65387
34	16	64.0	322101	10	ADA58431
35	16	64.0	322101	12	AD135046
36	15	60.0	150	6	ABN16840
37	15	60.0	336	8	ABX51125
38	15	60.0	480	6	ABT10253
39	15	60.0	525	10	ADD16647
40	15	60.0	529	4	AAK56957
41	15	60.0	529	4	AAK56957
42	15	60.0	535	12	ACH71148
43	15	60.0	572	3	AA87766
44	15	60.0	572	5	AA64048
45	15	60.0	572	12	ADP18786
46	15	60.0	593	5	ABV55376
47	15	60.0	703	2	AAZ15924
48	15	60.0	766	4	AAH34697
49	15	60.0	854	2	AAZ40809
50	15	60.0	854	11	ADY77786
51	15	60.0	854	12	ADP19107
52	15	60.0	871	10	AAH03505
53	15	60.0	996	8	ACA27493
54	15	60.0	1166	4	AA162903
55	15	60.0	1901	4	AAH17036
56	15	60.0	2000	6	ABZ15483
57	15	60.0	2136	4	AA162899
58	15	60.0	2581	10	ADFR9810
59	15	60.0	2988	2	AAZ0289
60	15	60.0	3001	3	AAH51775
61	15	60.0	3193	4	ABL04126
62	15	60.0	3380	4	ABL16198
63	15	60.0	3519	8	ACA33811
64	15	60.0	3552	10	ADC33399
65	15	60.0	3937	10	ADRS3913
66	15	60.0	4338	4	ABL17064
67	15	60.0	4785	5	AA572158
68	15	60.0	5064	2	AAV69110
69	15	60.0	5276	6	ABL22151
70	15	60.0	5617	4	ABL06156
71	15	60.0	5923	4	AA690320
72	15	60.0	6107	6	ABL70354
73	15	60.0	6107	6	ABL61315
74	15	60.0	6419	6	ABL52267
75	15	60.0	6760	3	AAZ58293
76	15	60.0	7546	4	ABL1880
77	15	60.0	8622	6	ABL14142
78	15	60.0	11021	4	ABL11102
79	15	60.0	11157	4	ABL09906
80	15	60.0	11515	4	ABL09906
81	15	60.0	13549	4	ABL07964
82	15	60.0	18624	6	ABL13703
83	15	60.0	22875	4	ABL18664
84	15	60.0	23439	4	ABL19006
85	15	60.0	3472	4	ABL19792
86	15	60.0	48000	9	AA677996
87	15	60.0	68230	9	ADA66349
88	15	60.0	68233	9	ADA03065
89	15	60.0	68233	11	ADA72803
90	15	60.0	68233	11	ADA727143
91	15	60.0	110000	6	ABA80521
92	15	60.0	110000	6	ABA80521
93	15	60.0	110000	10	ADP77343
94	15	60.0	110000	10	AA52246

ABL21530	Drosophila
ABL60141	Antirrhin
ABL32765	Human imm
ABL31271	Signal tr
ABL70232	Chemical
ABL61179	Human gen
ADA1645	Human sec
ACC50940	Human sec
ADD38155	CDNA clon
ADA57777	BAC fragm
ACR67367	Continuation (52 o
ACR65387	Continuation (3 of
ADA58431	Human PEO
AD135046	Human pro
ABN16840	Human ORF
ABX51125	Bovine BS
ABT10253	Human bre
ADD16647	DNA (Seg1
AAK56957	Human imm
AAK56957	Human CDN
ACH71148	Human sec
AA87766	Human sec
AA64048	CDNA enco
ADP18786	Human sec
ABV55376	Human pro
AAZ15924	Human gen
AAH34697	Human col
AAZ40809	Secreted
ADY77786	Human CDN
ADP19107	Human sec
AAH03505	Human CDN
ACA27493	Prokaryot
AA162903	Human gen
AAH17036	Human CDN
ABZ15483	Arabidops
AA162899	Human gen
ADFR9810	Tritepern
AAZ0289	Borrelia
AAH51775	Chromosom
ABL04126	Drosophila
ABL16198	Drosophila
ACA33811	Prokaryot
ADC33399	E. faeciu
ADRS3913	Human pro
ABL17064	Drosophila
AA572158	DNA enco
AAV69110	Neisseria
ABL22151	Human imm
ABL06156	Human imm
AA690320	Pea auxin
ABL70354	Chemical
ABL61315	Human gen
ABL52267	Human imm
AAZ58293	Sorghum r
ABL1880	Drosophila
ABL14142	Human imm
ABL11102	Drosophila
ABL09906	Human imm
ABL09906	Drosophila
ABL07964	Drosophila
ABL13703	Human imm
ABL18664	Drosophila
ABL19006	Drosophila
ABL19792	Drosophila
AA677996	Human cal
ADA66349	Mouse PEP
ADA03065	Mouse PEP
ADA72803	Mouse PEP
ADA727143	Mouse gen
ABA80521	Continuation (16 o
ABA80521	Continuation (20 o
ADP77343	Continuation (10 o
AA52246	Continuation (2 of

95	15	60.0	112132	6	ABK90888	ADK90888 Human ATP
96	15	60.0	112132	11	ADM56267	ADM56267 Human ATP
97	15	60.0	127197	5	AA161370	AA161370 Soybean 5
98	15	60.0	177851	8	AA157272	AA157272 BA438823-
99	15	60.0	266145	10	AD87477	AD87477 FOW1pox v
100	15	60.0	340449	8	AA152198	AA152198 Human sec

ALIGNMENTS

RESULT 1
AA161370 standard: DNA; 25 BP.

AC AA161370;

DT 02-JUL-2001 (first entry)

DE Lactobacillus brevis 23S rRNA/5S rRNA specific probe SEQ ID 73.

XX 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

XX Lactobacillus brevis.

XX DE19945964-A1.

PD 05-APR-2001.

PF 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PA (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K;

DR WPI; 2001-246136/26.

PT Detecting contaminating microorganisms in brewing, by nucleic acid
amplification and hybridization, either non-specific or genus- or species
specific.

PS Claim 9(1); Page 17; 48pp; German.

XX This invention describes a novel method for detecting microorganisms (A)
of importance in brewing which comprises treating a sample with at least
two primers (P1) that hybridize to a consensus region in the nucleic acid
of (A), at least part of the microbial nucleic acid is amplified, the
amplicon is treated with at least one probe (P2) that hybridizes
specifically with a sequence common to all (A) or specific for one or
more families, genera or species, and any formation of hybrids is
detected. The method is used to detect, identify and/or characterize
microorganisms in beer or brewing materials, particularly for detecting
contamination. The method may detect the entire range of contaminating
microbes, either as a general test for contamination or as a test
specific for particular genera or (sub)species. It is quicker than known
microbiological methods, and can detect several organisms in the same
sample, including organisms not presently recognized as contaminants. The
method provides an early indication of contamination and can be automated
for high throughput analysis

SO Sequence 25 BP; 11 A; 2 C; 4 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATATCTAG 25
DB 1 TCGAGATATATGATATATCTAG 25

RESULT 2
AA161566 standard: DNA; 267 BP.

XX AA161566;

DT 02-JUL-2001 (first entry)

DE Lactobacillus brevis 23S rRNA-spacer-5S rRNA DNA fragment.

XX 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

XX Lactobacillus brevis.

XX DE19945964-A1.

PD 05-APR-2001.

PF 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PA (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K;

DR WPI; 2001-246136/26.

PT Detecting contaminating microorganisms in brewing, by nucleic acid
amplification and hybridization, either non-specific or genus- or species
specific.

PS Claim 9(1); Page 9; 48pp; German.

XX This invention describes a novel method for detecting microorganisms (A)
of importance in brewing which comprises treating a sample with at least
two primers (P1) that hybridize to a consensus region in the nucleic acid
of (A), at least part of the microbial nucleic acid is amplified, the
amplicon is treated with at least one probe (P2) that hybridizes
specifically with a sequence common to all (A) or specific for one or
more families, genera or species, and any formation of hybrids is
detected. The method is used to detect, identify and/or characterize
microorganisms in beer or brewing materials, particularly for detecting
contamination. The method may detect the entire range of contaminating
microbes, either as a general test for contamination or as a test
specific for particular genera or (sub)species. It is quicker than known
microbiological methods, and can detect several organisms in the same
sample, including organisms not presently recognized as contaminants. The
method provides an early indication of contamination and can be automated
for high throughput analysis

SO Sequence 267 BP; 77 A; 51 C; 77 G; 62 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 5; Length 267;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCTAG 25
DB 125 TCGAGATATATGATATATCTAG 149

RESULT 3
AA13651/c
AA13651 standard: DNA; 1190 BP.

XX AA13651;

DT 19-MAR-1999 (first entry)

DE Enterococcus faecalis genome contig SEQ ID NO:714.

XX Enterococcus faecalis; contig; detection; Enterococcal infection;

Tue Nov 16 16:09:52 2004

us-10-088-666-21.olig10.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:28:49 ; Search time 36.1351 Seconds
(without alignments)
3138.794 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21 ccaagtcacacacgtagctgtc 21

Scoring table: OLIGO_NUC;
Gapop 60.0 ; Gapext 60.0

Searched: 3625171 seqs, 2700493622 residues

Word size: 10

Total number of hits satisfying chosen parameters: 28758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Listing first 100 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	76.2	365	16 US-10-424-599-100843	Sequence 100843, A
2	15	71.4	395	16 US-10-424-599-97864	Sequence 97864, A
3	15	71.4	621	17 US-10-767-795-4222	Sequence 4222, Ap
4	15	71.4	741	15 US-10-369-493-30596	Sequence 30596, A
5	15	71.4	1134	16 US-10-425-114-20775	Sequence 20775, A
6	15	71.4	1134	18 US-10-425-115-16878	Sequence 16878, A
7	15	71.4	1156	16 US-10-425-114-10548	Sequence 10548, A
8	15	71.4	1161	16 US-10-424-599-49681	Sequence 49681, A
9	15	71.4	1188	16 US-10-424-599-133586	Sequence 133586, A
10	15	71.4	168276	13 US-10-087-192-142	Sequence 142, App
11	14	66.7	254	17 US-10-437-963-8513	Sequence 8513, Ap
12	14	66.7	559	13 US-10-027-632-241832	Sequence 241832, Ap

C 13	14	66.7	559	15	US-10-027-632-241832	Sequence 241832, A
C 14	14	66.7	561	17	US-10-437-963-32347	Sequence 32347, A
C 15	14	66.7	591	15	US-10-029-386-7736	Sequence 7736, Ap
C 16	14	66.7	601	13	US-10-027-632-199748	Sequence 199748, A
C 17	14	66.7	601	13	US-10-027-632-199749	Sequence 199749, A
C 18	14	66.7	601	13	US-10-027-632-199748	Sequence 199748, A
C 19	14	66.7	601	15	US-10-027-632-199749	Sequence 199749, A
C 20	14	66.7	1203	13	US-10-027-632-199747	Sequence 199747, A
C 21	14	66.7	1203	15	US-10-027-632-199747	Sequence 199747, A
C 22	14	66.7	1613	17	US-10-451-467A-259	Sequence 259, App
C 23	14	66.7	1650	16	US-10-282-1222-30915	Sequence 30915, A
C 24	14	66.7	1820	17	US-10-437-963-97628	Sequence 97628, A
C 25	14	66.7	2000	16	US-10-260-238-1774	Sequence 1774, Ap
C 26	14	66.7	2347	16	US-10-424-599-54300	Sequence 54300, A
C 27	14	66.7	2732	15	US-10-369-493-36202	Sequence 36202, A
C 28	14	66.7	2796	17	US-10-437-963-97626	Sequence 97626, A
C 29	14	66.7	3825	9	US-09-746-491-11	Sequence 11, App1
C 30	14	66.7	19639	8	US-08-910-386A-6	Sequence 6, App1
C 31	14	66.7	143973	13	US-10-087-192-442	Sequence 442, App
C 32	13	61.9	25	15	US-10-098-263B-47079	Sequence 47079, A
C 33	13	61.9	180	17	US-10-767-701-23303	Sequence 23203, A
C 34	13	61.9	207	9	US-09-960-352-4229	Sequence 4229, Ap
C 35	13	61.9	211	9	US-09-960-352-4061	Sequence 2061, Ap
C 36	13	61.9	229	9	US-09-960-352-4220	Sequence 4220, Ap
C 37	13	61.9	239	9	US-09-960-352-13451	Sequence 13451, A
C 38	13	61.9	242	9	US-09-960-352-2998	Sequence 2998, Ap
C 39	13	61.9	247	9	US-09-960-352-183	Sequence 183, App
C 40	13	61.9	281	9	US-09-960-352-9850	Sequence 9850, Ap
C 41	13	61.9	302	9	US-09-960-352-7737	Sequence 7737, Ap
C 42	13	61.9	391	9	US-09-960-352-86	Sequence 86, App1
C 43	13	61.9	308	9	US-09-960-352-10392	Sequence 10392, A
C 44	13	61.9	310	16	US-10-424-599-104720	Sequence 104720, A
C 45	13	61.9	324	10	US-09-764-891-692	Sequence 692, App
C 46	13	61.9	325	17	US-10-091-572-56	Sequence 56, App1
C 47	13	61.9	325	14	US-10-437-963-33253	Sequence 33253, A
C 48	13	61.9	337	10	US-09-764-891-6697	Sequence 6697, Ap
C 49	13	61.9	337	10	US-09-764-891-6698	Sequence 6698, Ap
C 50	13	61.9	337	14	US-10-091-572-558	Sequence 558, App
C 51	13	61.9	337	14	US-10-091-572-559	Sequence 559, App
C 52	13	61.9	340	9	US-09-960-352-2761	Sequence 2761, App
C 53	13	61.9	341	9	US-09-960-352-2191	Sequence 2191, Ap
C 54	13	61.9	341	16	US-10-424-599-12628	Sequence 12628, A
C 55	13	61.9	342	9	US-09-960-352-11482	Sequence 11482, A
C 56	13	61.9	348	9	US-09-960-352-9403	Sequence 9403, Ap
C 57	13	61.9	358	9	US-09-960-352-8472	Sequence 8472, Ap
C 58	13	61.9	360	9	US-09-960-352-2079	Sequence 2079, Ap
C 59	13	61.9	360	9	US-09-960-352-12857	Sequence 12857, A
C 60	13	61.9	366	9	US-09-960-352-8659	Sequence 8659, Ap
C 61	13	61.9	366	9	US-09-960-352-14094	Sequence 14094, A
C 62	13	61.9	378	9	US-09-960-352-4223	Sequence 4223, Ap
C 63	13	61.9	378	9	US-09-960-352-2966	Sequence 2966, Ap
C 64	13	61.9	384	9	US-09-960-352-215	Sequence 215, App
C 65	13	61.9	384	17	US-10-767-701-30666	Sequence 30666, A
C 66	13	61.9	385	9	US-09-960-352-10641	Sequence 10641, A
C 67	13	61.9	387	9	US-09-960-352-13175	Sequence 13175, A
C 68	13	61.9	391	9	US-09-783-590-9798	Sequence 9798, Ap
C 69	13	61.9	391	9	US-09-960-352-10595	Sequence 10595, A
C 70	13	61.9	395	9	US-09-960-352-2505	Sequence 2505, Ap
C 71	13	61.9	398	9	US-09-960-352-7013	Sequence 7013, Ap
C 72	13	61.9	399	9	US-09-960-352-9396	Sequence 9396, Ap
C 73	13	61.9	404	9	US-09-960-352-1179	Sequence 1179, App
C 74	13	61.9	404	9	US-09-960-352-2789	Sequence 2789, Ap
C 75	13	61.9	404	9	US-09-960-352-9795	Sequence 9795, Ap
C 76	13	61.9	405	9	US-09-960-352-7246	Sequence 7246, Ap
C 77	13	61.9	405	10	US-09-803-719-766	Sequence 766, App
C 78	13	61.9	405	16	US-10-424-599-45029	Sequence 45029, A
C 79	13	61.9	406	9	US-09-960-352-12380	Sequence 12380, A
C 80	13	61.9	407	9	US-09-960-352-9848	Sequence 9848, Ap
C 81	13	61.9	416	9	US-09-960-352-11517	Sequence 11517, A
C 82	13	61.9	417	9	US-09-960-352-10053	Sequence 10053, A
C 83	13	61.9	417	16	US-10-424-599-124751	Sequence 124751, A
C 84	13	61.9	419	9	US-09-960-352-7509	Sequence 7509, App
C 85	13	61.9	420	9	US-09-960-352-6117	Sequence 6117, Ap

86 13 61.9 427 9 US-09-960-352-7803 Sequence 7803, Ap
 87 13 61.9 430 9 US-09-960-352-5911 Sequence 5911, Ap
 88 13 61.9 441 17 US-10-437-963-25328 Sequence 25328, A
 89 13 61.9 458 17 US-09-960-352-9145 Sequence 9145, Ap
 90 13 61.9 468 17 US-10-437-963-18379 Sequence 18379, A
 91 13 61.9 496 15 US-10-027-632-126514 Sequence 126514, A
 92 13 61.9 496 15 US-10-027-632-126514 Sequence 126514, A
 93 13 61.9 499 9 US-09-864-761-10086 Sequence 10086, A
 94 13 61.9 505 13 US-10-027-632-199383 Sequence 199383, A
 95 13 61.9 508 15 US-10-027-632-199383 Sequence 199383, A
 96 13 61.9 505 16 US-10-424-599-18750 Sequence 18750, A
 97 13 61.9 517 16 US-10-424-599-24383 Sequence 24383, A
 98 13 61.9 526 16 US-10-424-599-41897 Sequence 41897, A
 99 13 61.9 535 9 US-09-867-550-1313 Sequence 1313, Ap
 100 13 61.9 542 13 US-10-027-632-50508 Sequence 50508, A

ALIGNMENTS

RESULT 1
 US-10-424-599-100843
 ; Sequence 100843, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 100843
 ; LENGTH: 365
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_62074C.1
 US-10-424-599-100843

Query Match 76.2%; Score 16; DB 16; Length 365;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGTCAACAAGTACTT 19
 |||||
 DB 124 AGTCAACAAGTACTT 139

RESULT 2
 US-10-424-599-97864/C
 ; Sequence 97864, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 97864
 ; LENGTH: 395
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_59385C.1

US-10-424-599-97864
 Query Match 71.4%; Score 15; DB 16; Length 395;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GTCAACAAGTACTT 19
 |||||
 DB 125 GTCAACAAGTACTT 111

RESULT 3
 US-10-767-795-4222/C
 ; Sequence 4222, Application US/10767795
 ; Publication No. US2004018130A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21(53534)B
 ; CURRENT APPLICATION NUMBER: US/10/767,795
 ; CURRENT FILING DATE: 2004-01-30
 ; NUMBER OF SEQ ID NOS: 117596
 ; SEQ ID NO 4222
 ; LENGTH: 621
 ; TYPE: DNA
 ; ORGANISM: Gossypium hirsutum
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C111_9
 US-10-767-795-4222

Query Match 71.4%; Score 15; DB 17; Length 621;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAGTCAACAAGTACTG 17
 |||||
 DB 87 AAGTCAACAAGTACTG 73

RESULT 4
 US-10-369-493-30596
 ; Sequence 30596, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 30596
 ; LENGTH: 741
 ; TYPE: DNA
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-30596

Query Match 71.4%; Score 15; DB 15; Length 741;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGTCAACAAGCT 15
 |||||
 DB 88 CCAAGTCAACAAGCT 102

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 14:18:54 ; Search time 8.95496 Seconds
(without alignments)
1666.850 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21

Sequence: 1 ccaagtcacacagctgtgt 21

Scoring table: OLIGO-NUC;
Gapop 60.0, Gapext 60.0

Searched: 824507 seqs, 355394441 residues

Word size: 10

Total number of hits satisfying chosen parameters: 4843

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	66.7	610	4	US-09-270-767-12339
2	13	61.9	186	4	US-09-248-796A-7468
3	13	61.9	189	4	US-09-583-110-2597
4	13	61.9	360	4	US-09-134-000C-796
5	13	61.9	396	4	US-09-252-991A-6239
6	13	61.9	525	4	US-09-248-796A-1527
7	13	61.9	564	4	US-09-583-110-2598
8	13	61.9	639	4	US-09-248-796A-6195
9	13	61.9	708	1	US-07-794-398B-1
10	13	61.9	708	1	US-08-002-866-1
11	13	61.9	780	4	US-09-248-796A-6461
12	13	61.9	954	4	US-09-248-796A-2947
13	13	61.9	963	4	US-09-328-352-3811
14	13	61.9	997	4	US-09-690-454-33
15	13	61.9	1089	4	US-09-291-299A-5
16	13	61.9	1101	4	US-09-248-796A-1389
17	13	61.9	1236	4	US-09-016-434-717
18	13	61.9	1243	1	US-08-702-344-27
19	13	61.9	1266	4	US-09-248-796A-314
20	13	61.9	1272	4	US-09-489-039A-5398
21	13	61.9	1295	4	US-09-219-194-3
22	13	61.9	1399	4	US-09-248-796A-5857
23	13	61.9	1392	4	US-09-270-767-10521
24	13	61.9	1311	4	US-09-248-796A-5537
25	13	61.9	1317	4	US-09-248-796A-5750
26	13	61.9	1318	4	US-09-270-767-25088
27	13	61.9	1356	4	US-09-489-039A-5355
28	13	61.9	1413	4	US-09-134-000C-2063
29	13	61.9	1505	4	US-08-291-299A-2
30	13	61.9	1523	4	US-09-291-299A-4
31	13	61.9	1950	4	US-09-270-767-13177
32	13	61.9	2073	4	US-09-248-796A-5066
33	13	61.9	2130	4	US-09-583-110-2452
34	13	61.9	2133	4	US-09-248-796A-302
35	13	61.9	2218	2	US-08-421-044-1
36	13	61.9	2307	4	US-08-792-924-43
37	13	61.9	2342	1	US-08-441-700-1
38	13	61.9	2669	4	US-08-118-101A-3
39	13	61.9	2670	4	US-09-614-221A-540
40	13	61.9	4565	3	US-08-776-265-2
41	13	61.9	4565	4	US-09-398-184-2
42	13	61.9	5240	4	US-09-171-337A-2
43	13	61.9	5240	4	US-09-631-022-2
44	13	61.9	8791	5	PCT-US96-01735-5
45	13	61.9	8791	5	US-09-566-921-45
46	13	61.9	9848	4	US-09-385-222A-3
47	13	61.9	15363	4	US-08-961-527-139
48	12	57.1	85	3	US-09-364-543-57
49	12	57.1	90	3	US-08-974-549A-685
50	12	57.1	90	3	US-08-974-549A-686
51	12	57.1	90	3	US-09-721-456-685
52	12	57.1	90	4	US-09-721-456-686
53	12	57.1	108	5	PCT-US95-09589-6
54	12	57.1	108	5	PCT-US95-09589A-6
55	12	57.1	227	4	US-09-513-899C-30518
56	12	57.1	267	4	US-09-252-911A-6214
57	12	57.1	286	4	US-09-270-767-14783
58	12	57.1	291	4	US-09-313-294A-6338
59	12	57.1	324	4	US-09-270-767-10980
60	12	57.1	330	5	PCT-US95-09589-1
61	12	57.1	330	5	PCT-US95-09589A-1
62	12	57.1	350	2	US-08-967-101-35
63	12	57.1	350	2	US-08-967-101-35
64	12	57.1	350	3	US-09-124-698-35
65	12	57.1	350	3	US-09-124-698-35
66	12	57.1	350	3	US-09-124-698-35
67	12	57.1	350	3	US-09-124-698-35
68	12	57.1	350	3	US-09-124-698-35
69	12	57.1	350	4	US-09-636-796A-35
70	12	57.1	350	4	US-08-431-048-35
71	12	57.1	358	4	US-09-016-434-707
72	12	57.1	372	4	US-09-134-000C-1089
73	12	57.1	400	4	US-08-781-986A-3674
74	12	57.1	400	4	US-08-781-986A-3674
75	12	57.1	442	4	US-09-248-796A-3904
76	12	57.1	444	4	US-09-248-796A-8988
77	12	57.1	454	4	US-09-221-017B-119
78	12	57.1	457	4	US-09-270-767-10199
79	12	57.1	461	4	US-09-513-999C-4086
80	12	57.1	462	4	US-09-270-767-10077
81	12	57.1	551	4	US-09-187-999-32
82	12	57.1	555	4	US-09-270-767-26085
83	12	57.1	554	4	US-09-543-681A-2575
84	12	57.1	570	3	US-08-998-816-159
85	12	57.1	600	4	US-09-325-932A-6
86	12	57.1	615	4	US-09-248-796A-5523
87	12	57.1	618	4	US-09-443-681A-3028
88	12	57.1	634	3	US-09-415-522-1
89	12	57.1	636	4	US-09-669-991-12
90	12	57.1	665	4	US-09-352-991A-6331
91	12	57.1	668	4	US-08-956-711E-3658
92	12	57.1	668	4	US-08-956-711E-3658
93	12	57.1	735	4	US-09-328-352-1471
94	12	57.1	749	4	US-09-270-767-13258
95	12	57.1	778	4	US-09-325-932A-7
96	12	57.1	783	4	US-09-352-991A-6415
97	12	57.1	819	2	US-08-441-857-3
98	12	57.1	819	3	US-08-642-807A-31
99	12	57.1	819	3	US-08-193-159-3
100	12	57.1	819	4	US-09-252-991A-6136

ALIGNMENTS

RESULT 1

US-09-270-767-12339
; Sequence 12339, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12339
; LENGTH: 610
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-12339

Query Match 66.7%; Score 14; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GTCAACAAGCTAGT 18
 |||||
DB 498 GTCAACAAGCTAGT 511

RESULT 2

US-09-248-796A-7468
; Sequence 7468, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7468
; LENGTH: 186
; TYPE: DNA
; ORGANISM: *Candida albicans*
US-09-248-796A-7468

Query Match 61.9%; Score 13; DB 4; Length 186;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGTCAACAC 13
 |||||
DB 4 CCAAGTCAACAC 16

RESULT 3

US-09-583-110-2597/C
; Sequence 2597, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to *Streptococcus*
; FILE REFERENCE: PAT00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2597
; LENGTH: 189
; TYPE: DNA
; ORGANISM: *Streptococcus pneumoniae*
US-09-583-110-2597

Query Match 61.9%; Score 13; DB 4; Length 189;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGTCAACAC 13
 |||||
DB 174 CCAAGTCAACAC 162

RESULT 4

US-09-134-000C-796/C
; Sequence 796, Application US/09134000C
; Patent No. 6617136
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 796
; LENGTH: 360
; TYPE: DNA
; ORGANISM: *Enterococcus faecalis*
US-09-134-000C-796

Query Match 61.9%; Score 13; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGTCAACAC 13
 |||||
DB 265 CCAAGTCAACAC 253

RESULT 5

US-09-252-991A-6239
; Sequence 6239, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6239
; LENGTH: 396
; TYPE: DNA
; ORGANISM: *Pseudomonas aeruginosa*

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:59:59 ; Search time 40.6757 Seconds
(without alignments)
2710.166 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21

Sequence: 1 ccaagtcacacagtagtgc 21

Scoring table: Oligo NUC
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size: 10

Total number of hits satisfying chosen parameters: 31333

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : N_Geneseq_23sep04:*

1: geneseqn1980s:*\n2: geneseqn1980s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	5	AA61586 Lactobaci
2	21	100.0	267	5	AA61586 Lactobaci
3	18	85.7	2215	3	AA243884 C. vicina
4	14	66.7	546	6	ABN23311 Human ORF
5	14	66.7	591	12	ACH74541 Aspergilli
6	14	66.7	668	3	AFR13909 Aspergilli
7	14	66.7	729	6	AB862906 Selected
8	14	66.7	729	6	ABT11511 Yeast sel
9	14	66.7	750	11	ADM45774 Insect re
10	14	66.7	766	6	AB862991 Selected
11	14	66.7	766	6	ABT11596 Yeast sel
12	14	66.7	789	10	ADK54602 Plant DNA
13	14	66.7	789	11	ADM45503 Insect re
14	14	66.7	812	6	AB863068 Selected
15	14	66.7	813	6	ABT11673 DNA Seq I
16	14	66.7	935	10	ADCT2643 DNA Seq I
17	14	66.7	935	10	ADCT2635 DNA Seq I
18	14	66.7	1613	6	ABQ76417 S. cerevi
19	14	66.7	1650	4	ACA43045 Prokaryot
20	14	66.7	1680	4	ABT21577 Drosophili
21	14	66.7	2000	12	ADJ40734 Plant CDN

22	14	66.7	3070	10	ADD47979 Human gen
23	14	66.7	3591	12	ADN13951 Human pro
24	14	66.7	3825	4	AAC86158 FCIR6 nuc
25	14	66.7	4141	4	ABT21576 Drosophili
26	14	66.7	19639	2	AAK23524 O. longis
27	14	66.7	110000	12	ADN46845_10 Continuation (11 o
28	14	66.7	110000	12	ADN47591_10 Continuation (11 o
29	14	66.7	110000	12	ADN46123_10 Continuation (11 o
30	14	66.7	110000	12	ADN47209_10 Continuation (11 o
31	14	66.7	110000	12	ADN46464_10 Continuation (11 o
32	14	66.7	110000	12	ADN47960_10 Continuation (11 o
33	13	61.9	25	9	ACT47088 Human mic
34	13	61.9	102	6	ABN71387 Streptoco
35	13	61.9	121	12	ADK33359 Polynucle
36	13	61.9	207	8	ABX39064 Bx3696 Bx3696
37	13	61.9	211	8	ABX36896 Bx3696 Bx3696
38	13	61.9	229	8	ABX39055 Bx3696 Bx3696
39	13	61.9	239	8	ABX46286 Bx3696 Bx3696
40	13	61.9	242	8	ABX37833 Bx3696 Bx3696
41	13	61.9	244	6	AB862927 Selected
42	13	61.9	247	8	ABT11532 Yeast sel
43	13	61.9	248	5	ABX35018 Bx3696 Bx3696
44	13	61.9	248	5	AAI69939 Bx3696 Bx3696
45	13	61.9	281	8	ABX44685 Bx3696 Bx3696
46	13	61.9	292	8	ABX42572 Bx3696 Bx3696
47	13	61.9	301	8	ABX34921 Bx3696 Bx3696
48	13	61.9	308	8	ABX45227 Bx3696 Bx3696
49	13	61.9	324	4	AAI00691 Human rep
50	13	61.9	324	5	AAI40115 DNA encod
51	13	61.9	337	4	AAI04009 Human rep
52	13	61.9	337	4	AAI04010 Human rep
53	13	61.9	337	5	AAI40406 DNA encod
54	13	61.9	337	5	AAI40407 DNA encod
55	13	61.9	340	8	ABX37596 Bx3696 Bx3696
56	13	61.9	341	8	ABX37026 Bx3696 Bx3696
57	13	61.9	342	8	ABX46317 Bx3696 Bx3696
58	13	61.9	345	10	ADK59305 Plant DNA
59	13	61.9	348	8	ABX44238 Bx3696 Bx3696
60	13	61.9	352	3	AAH30860 Human col
61	13	61.9	358	8	ABX43307 Bx3696 Bx3696
62	13	61.9	359	3	AAA43856 Human sec
63	13	61.9	360	8	ABX36914 Bx3696 Bx3696
64	13	61.9	360	8	ABX47692 Bx3696 Bx3696
65	13	61.9	360	10	ADH82921 Enterococ
66	13	61.9	363	5	ADM19610 Novel hum
67	13	61.9	366	8	ABX43494 Bx3696 Bx3696
68	13	61.9	366	8	ABX48929 Bx3696 Bx3696
69	13	61.9	368	8	ABX39058 Bx3696 Bx3696
70	13	61.9	378	8	ABX37801 Bx3696 Bx3696
71	13	61.9	382	8	ABX35050 Bx3696 Bx3696
72	13	61.9	385	8	ABX45476 Bx3696 Bx3696
73	13	61.9	387	8	ABX48010 Bx3696 Bx3696
74	13	61.9	391	8	ABX45430 Bx3696 Bx3696
75	13	61.9	395	8	ABX37340 Bx3696 Bx3696
76	13	61.9	396	11	ABD07635 Pseudomon
77	13	61.9	398	11	ABD07635 Pseudomon
78	13	61.9	398	8	ABX41848 Bx3696 Bx3696
79	13	61.9	399	8	ABX44231 Bx3696 Bx3696
80	13	61.9	404	8	ABX44724 Bx3696 Bx3696
81	13	61.9	404	8	ABX35014 Bx3696 Bx3696
82	13	61.9	404	8	ABX37630 Bx3696 Bx3696
83	13	61.9	405	3	AAI51811 Arabidops
84	13	61.9	405	4	AAI51811 Arabidops
85	13	61.9	405	8	ABX37708 Novel hum
86	13	61.9	406	8	ABX42081 Bx3696 Bx3696
87	13	61.9	407	8	ABX47215 Bx3696 Bx3696
88	13	61.9	408	3	AAI4681 Bx3696 Bx3696
89	13	61.9	416	8	ABX46352 Bx3696 Bx3696
90	13	61.9	417	8	ABX44888 Bx3696 Bx3696
91	13	61.9	419	8	ABX42344 Bx3696 Bx3696
92	13	61.9	420	8	ABX40952 Bx3696 Bx3696
93	13	61.9	427	8	ABX42638 Bx3696 Bx3696
94	13	61.9	430	8	ABX40746 Bx3696 Bx3696

95	13	61.9	444	6	ABN68526	Abn68526 Streptococcus
96	13	61.9	455	9	AC117009	AC117009 DNA clone
97	13	61.9	456	8	AB252858	AB252858 Aspergillus
98	13	61.9	458	8	ABX43980	ABX43980 Bovine ES
99	13	61.9	489	10	ADG37805	ADG37805 Aspergillus
100	13	61.9	499	4	AA133114	AA133114 Probe #18

ALIGNMENTS

RESULT 1

AAAF61586 standard; DNA; 21 BP.

ID AAAF61586;

AC AAAF61586;

DE Lactobacillus brevis 23S rRNA/5S rRNA specific DNA probe.

KM 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

OS Lactobacillus brevis.

PN DE19945964-A1.

PD 05-APR-2001.

PF 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PS (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K;

DR WPI; 2001-246136/26.

PT Detecting contaminating microorganisms in brewing, by nucleic acid

PT amplification and hybridization, either non-specific or genus- or species

PT -specific.

PS Claim 9(1); Page 15; 48pp; German.

CC This invention describes a novel method for detecting microorganisms (A)

CC of importance in brewing which comprises treating a sample with at least

CC two primers (P1) that hybridize to a consensus region in the nucleic acid

CC of (A), at least part of the microbial nucleic acid is amplified, the

CC amplicon is treated with at least one probe (P2) that hybridizes

CC specifically with a sequence common to all (A) or specific for one or

CC more families, genera or species, and any formation of hybrids is

CC detected. The method is used to detect, identify and/or characterize

CC microorganisms in beer or brewing materials, particularly for detecting

CC contamination. The method may detect the entire range of contaminating

CC microbes, either as a general test for contamination or as a test

CC specific for particular genera or (sub)species. It is quicker than known

CC microbiological methods, and can detect several organisms in the same

CC sample, including organisms not presently recognized as contaminants. The

CC method provides an early indication of contamination and can be automated

CC for high throughput analysis

SQ Sequence 21 BP; 7 A; 5 C; 4 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 5; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.024; Mismatches 0; Gaps 0;

Matches 21; Conservative 0; Indels 0; Gaps 0;

CY 1 CCAAGTCACACGTAAGTGT 21

DB 1 CCAAGTCACACGTAAGTGT 21

RESULT 2

ID AAAF61566 standard; DNA; 267 BP.

AC AAAF61566;

DE Lactobacillus brevis 23S rRNA-spacer-5S rRNA DNA fragment.

KM 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

OS Lactobacillus brevis.

PN DE19945964-A1.

PD 05-APR-2001.

PF 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PS (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K;

DR WPI; 2001-246136/26.

PT Detecting contaminating microorganisms in brewing, by nucleic acid

PT amplification and hybridization, either non-specific or genus- or species

PT -specific.

PS Claim 9(1); Page 9; 48pp; German.

CC This invention describes a novel method for detecting microorganisms (A)

CC of importance in brewing which comprises treating a sample with at least

CC two primers (P1) that hybridize to a consensus region in the nucleic acid

CC of (A), at least part of the microbial nucleic acid is amplified, the

CC amplicon is treated with at least one probe (P2) that hybridizes

CC specifically with a sequence common to all (A) or specific for one or

CC more families, genera or species, and any formation of hybrids is

CC detected. The method is used to detect, identify and/or characterize

CC microorganisms in beer or brewing materials, particularly for detecting

CC contamination. The method may detect the entire range of contaminating

CC microbes, either as a general test for contamination or as a test

CC specific for particular genera or (sub)species. It is quicker than known

CC microbiological methods, and can detect several organisms in the same

CC sample, including organisms not presently recognized as contaminants. The

CC method provides an early indication of contamination and can be automated

CC for high throughput analysis

SQ Sequence 267 BP; 77 A; 51 C; 77 G; 62 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 5; Length 267;

Best Local Similarity 100.0%; Pred. No. 0.024; Mismatches 0; Gaps 0;

Matches 21; Conservative 0; Indels 0; Gaps 0;

CY 1 CCAAGTCACACGTAAGTGT 21

DB 103 CCAAGTCACACGTAAGTGT 123

RESULT 3

ID AA243884 standard; cDNA; 2215 BP.

AC AA243884;

DE 10-MAR-2000 (first entry)

KM C. vicina LSP-2 cDNA.

KM Arylphorin; calliphorin; ABP; receptor-binding domain; plant protection;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 12:43:29 ; Search time 163.964 Seconds
(without alignments)
6056.719 Million cell updates/sec

Title: US-10-088-666-21
Perfect score: 21
Sequence: 1 ccaagtcacacgacgtatgtc 21

Scoring table: OLIGO-NGC
Gapex 60.0 , Gapext 60.0

Searched: 4526729 segs, 23644849745 residues
Word-size: 10

Total number of hits satisfying chosen parameters: 181462

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	6	AX105749 Sequence
2	21	100.0	267	6	AX105729 Sequence
3	18	85.7	2215	8	CV089789 Sequence
4	16	76.2	836	8	AF149254 Brachycephalus
5	16	76.2	6327	14	AF016914 Brachycephalus
6	16	76.2	6327	14	AF571289 Brachycephalus
7	16	76.2	38703	2	AC129901 Brachycephalus
8	16	76.2	40689	2	AC129903 Brachycephalus
9	16	76.2	107579	8	AC016162 Arabidopsis
10	16	76.2	110000	8	CR380949_3 Arabidopsis
11	16	76.2	110000	8	CR382130_27 Arabidopsis
12	16	76.2	160681	2	AC118586 Arabidopsis
13	16	76.2	170274	2	AC146726 Arabidopsis
14	16	76.2	198767	2	AC116847 Arabidopsis
15	16	76.2	228453	2	AC146729 Arabidopsis
16	16	76.2	254684	2	AC094805 Arabidopsis
17	16	76.2	261162	1	AE017257 Arabidopsis
18	16	76.2	268619	2	AC130037 Arabidopsis
19	16	76.2	268619	2	AC130037 Arabidopsis

20	16	76.2	296100	1	AP005086 Arabidopsis
21	15	71.4	393	8	NBJ228073 Arabidopsis
22	15	71.4	497	8	PAB296697 Arabidopsis
23	15	71.4	1390	3	AK114505 Arabidopsis
24	15	71.4	3375	3	AY216495 Arabidopsis
25	15	71.4	4407	14	HB6DNAFOL Arabidopsis
26	15	71.4	4549	3	CE1FHD1 Arabidopsis
27	15	71.4	10060	3	AY216496 Arabidopsis
28	15	71.4	18412	14	AY446863 Arabidopsis
29	15	71.4	18805	3	CE38B2 Arabidopsis
30	15	71.4	32945	2	AC149342 Arabidopsis
31	15	71.4	37400	5	AF022814 Arabidopsis
32	15	71.4	38491	14	HB61MEXRH Arabidopsis
33	15	71.4	68167	5	AL731656 Arabidopsis
34	15	71.4	72443	2	AC019973 Arabidopsis
35	15	71.4	74903	8	AC005312 Arabidopsis
36	15	71.4	75803	8	AB016890 Arabidopsis
37	15	71.4	98957	8	AC004136 Arabidopsis
38	15	71.4	105246	8	AC139747 Arabidopsis
39	15	71.4	112972	3	AC006470 Arabidopsis
40	15	71.4	11685	8	AC141923 Arabidopsis
41	15	71.4	117835	2	AC024397 Arabidopsis
42	15	71.4	135052	5	AL731859 Arabidopsis
43	15	71.4	140583	9	AC010978 Arabidopsis
44	15	71.4	147983	8	AC118981 Arabidopsis
45	15	71.4	151355	2	AC106821 Arabidopsis
46	15	71.4	151979	2	AC025113 Arabidopsis
47	15	71.4	159321	14	HB6VAGNM Arabidopsis
48	15	71.4	160518	2	AC149942 Arabidopsis
49	15	71.4	160566	8	AL840634 Arabidopsis
50	15	71.4	164805	8	OSJN00174 Arabidopsis
51	15	71.4	178563	5	AL928650 Arabidopsis
52	15	71.4	186444	3	AC093439 Arabidopsis
53	15	71.4	188843	2	AC139844 Arabidopsis
54	15	71.4	193505	2	CR628366 Arabidopsis
55	15	71.4	193723	2	AC107725 Arabidopsis
56	15	71.4	197658	2	AC124208 Arabidopsis
57	15	71.4	213436	2	AC139299 Arabidopsis
58	15	71.4	216920	2	EX649356 Arabidopsis
59	15	71.4	220226	3	AC110124 Arabidopsis
60	15	71.4	222787	3	AE003637 Arabidopsis
61	15	71.4	226343	5	AL954845 Arabidopsis
62	15	71.4	235582	2	AC097134 Arabidopsis
63	15	71.4	246770	2	AC096461 Arabidopsis
64	15	71.4	250022	2	AE014824 Arabidopsis
65	15	71.4	256938	2	AC097183 Arabidopsis
66	15	71.4	258149	1	AC098277 Arabidopsis
67	15	71.4	300861	1	AE016928 Arabidopsis
68	15	71.4	347400	1	AP003591 Arabidopsis
69	15	71.4	347400	1	AP003591 Arabidopsis
70	15	71.4	347400	1	AP003591 Arabidopsis
71	14	66.7	471	3	AF540396 Arabidopsis
72	14	66.7	546	6	CO449339 Arabidopsis
73	14	66.7	752	14	GL424350 Arabidopsis
74	14	66.7	787	14	OYVWVAJ57 Arabidopsis
75	14	66.7	787	14	NCURSAM Arabidopsis
76	14	66.7	811	14	CVAJ8588 Arabidopsis
77	14	66.7	811	14	CVAJ8588 Arabidopsis
78	14	66.7	811	14	CVAJ8587 Arabidopsis
79	14	66.7	811	14	CVAJ8587 Arabidopsis
80	14	66.7	811	14	CVAJ8587 Arabidopsis
81	14	66.7	811	14	CVAJ8587 Arabidopsis
82	14	66.7	811	14	CVAJ8587 Arabidopsis
83	14	66.7	811	14	CVAJ8587 Arabidopsis
84	14	66.7	811	14	CVAJ8587 Arabidopsis
85	14	66.7	811	14	CVAJ8587 Arabidopsis
86	14	66.7	811	14	CVAJ8587 Arabidopsis
87	14	66.7	811	14	CVAJ8587 Arabidopsis
88	14	66.7	811	14	CVAJ8587 Arabidopsis
89	14	66.7	811	14	CVAJ8587 Arabidopsis
90	14	66.7	811	14	CVAJ8587 Arabidopsis
91	14	66.7	811	14	CVAJ8587 Arabidopsis
92	14	66.7	811	14	CVAJ8587 Arabidopsis

c 93 14 66.7 2739 14 OYMWVAJ51 AJ002451 Oira yell
 c 94 14 66.7 2744 14 CLCVAJ452 AJ002452 Cotton 1e
 c 95 14 66.7 2744 14 CLCVAJ455 AJ002455 Cotton 1e
 c 96 14 66.7 2902 8 SCYKR076N Z28301 S.cerevisia
 c 97 14 66.7 2964 8 AY372822 Magnapor AY372822 Magnapor
 c 98 14 66.7 3168 5 AY247207 Gasterost AY247207 Gasterost
 c 99 14 66.7 3284 8 SCYLR067C Z73239 S.cerevisia
 100 14 66.7 3825 6 AX181654 AX181654 Sequence

ALIGNMENTS

RESULT 1
 AX105749 21 bp DNA linear PAT 30-APR-2001
 LOCUS Sequence 21 from Patent WO0123605.
 DEFINITION AX105749
 ACCESSION AX105749
 VERSION AX105749.1 GI:13921762
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES

1
 Fandke, M., Gasch, A. and Berghof, K.
 Method and nucleic acids for determining the presence of
 micro-organisms specific to the brewing process
 Patent: WO 0123605-A 21 05-APR-2001;
 Biorecon Diagnostics GmbH (DE)
 Location/Qualifiers
 1..21
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="spezifische Sequenz fur Lactobacillus brevis"

ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTCAACACGAGTGTGT 21
 1 CCAGTCAACACGAGTGTGT 21
 Db 1 CCAGTCAACACGAGTGTGT 21

RESULT 2
 AX105729 267 bp DNA linear PAT 30-APR-2001
 LOCUS Sequence 1 from Patent WO0123605.
 DEFINITION AX105729
 ACCESSION AX105729
 VERSION AX105729.1 GI:13921742
 KEYWORDS
 SOURCE
 ORGANISM

1
 Lactobacillus brevis
 Lactobacillus brevis
 Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 Lactobacillus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES

1
 Fandke, M., Gasch, A. and Berghof, K.
 Method and nucleic acids for determining the presence of
 micro-organisms specific to the brewing process
 Patent: WO 0123605-A 1 05-APR-2001;
 Biorecon Diagnostics GmbH (DE)
 Location/Qualifiers
 1..267
 /organism="Lactobacillus brevis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1580"

ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 267;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTCAACACGAGTGTGT 21
 1 CCAGTCAACACGAGTGTGT 123
 Db 103 CCAGTCAACACGAGTGTGT 123

RESULT 3
 CVU89789/c 2215 bp mRNA linear INV 24-JUN-1998
 LOCUS Calliphora vicina hexamerin LSP-2 mRNA, complete cds.
 DEFINITION U89789
 ACCESSION U89789
 VERSION U89789.1 GI:1890113
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED

2 (bases 1 to 2215)
 Scheller, K., Koelling, C., Schroer, B. and Burmeister, T.
 Direct Submission
 Submitted (17-FEB-1997) Biologie du Developpement, Institut Jacques
 Monod, 2 place Jussieu, Paris 75005, France
 Location/Qualifiers
 1..2215
 /organism="Calliphora vicina"
 /mol_type="mRNA"
 /db_xref="taxon:7373"
 16..2121
 /codon_start=1
 /product="hexamerin LSP-2"
 /protein_id="AAC24157.1"
 /db_xref="GI:1890114"
 /translation="MKSFLVAVAVIALAPALASSKHIESKVDKDFLVKQKMDILQ
 HYQDVFVKYDESIVTYKPMHEADVEIDIDKFELMWDHDIYDEIFSPYER
 HVEYAVHILRLPYFPAKDNSTPHAVYMAITVYNKOLFISTVASIHRDGLGIVPA
 IYBIQPMHFPDADLEIAKFRMGHNVKXLDINLVNTYKSNVGDVNYEISLS
 YPLEDIGNAFRTYINLDIPFMANGVGHEILKDRGEPFYLEFHQILSRVYLERLSV
 DGEIPHENIYQYERHGFNSLHYTKGVSFPDRDNHNFYTDNYQAKDIYVGARL
 AQYVDYTNDEYLVAVKLGNTIQGNDSVNFVKYSLDLHYELNVBEGPFGYKGEEL
 PALIMHETSLDPLFYISYKDIVSYVRLAKTYPEYKXDLFPGVKIDSVHMPBGL
 TTYFEYEDADISNAVVEVSSSSADALFKFGKNSQLHGOSFVIAKQRLRNHKEFE
 YTLDDVMSKSKRAIVKVFIPKPYDENGLIHEIDNYMNFELDHVYVLDLAGVNHILKR
 SSDDFTWVYSKPTVLEIYOKYMDATNSGKRPINODEAHGCVPGSIMLPEKKGGMF
 POFYIVYFPHAEVQYQSTYDPVYISCGIGSGSRFVDTLPFGFPFNRPVRHDIYFNV
 NIKFYDIKIYHMDHTNLV"

CDS

ORIGIN

Query Match 85.7%; Score 18; DB 3; Length 2215;
 Best Local Similarity 100.0%; Pred. No. 4.7;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTCAACACGAGTGT 18
 1 CCAGTCAACACGAGTGT 1670
 Db 1687 CCAGTCAACACGAGTGT 1670

RESULT 4
 AF149254 836 bp DNA linear PLN 06-JUN-2001
 LOCUS Brassica napus resistance-related marker.
 DEFINITION AF149254
 ACCESSION AF149254
 VERSION AF149254.1 GI:14317917
 KEYWORDS
 SOURCE
 ORGANISM

Brassica napus (rape)
 Brassica napus

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 21; DB 6; Length 267;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCAGTCAACAGTGTGT 21
103 CCAGTCAACAGTGTGT 123

RESULT 3
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

AY327391 4692 bp DNA linear VRT 23-JUL-2004
Basilieuterus rivularis NADH dehydrogenase subunit 2, cytochrome
oxidase subunit I, cytochrome oxidase subunit II, ATP-synthase 8,
and ATP-synthase 6 genes, complete cds; mitochondrial.
AY327391.1 GI:39979797
mitochondrion Basilieuterus rivularis (neotropical river warbler)
Basilieuterus rivularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
Fringillidae; Emberizinae; Basilieuterus.
1 (bases 1 to 4692)
Lovelte, J. J.
Molecular phylogeny and plumage signal evolution in a trans Andean
and circum Amazonian avian species complex
Mol. Phylogenet. Evol. 32 (2), 512-523 (2004)
2 (bases 1 to 4692)
Lovelte, J. J.
Direct Submission
Submitted (20-JUN-2003) Laboratory of Ornithology, Cornell
University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA
Location/Qualifiers
1..4692
/organism="Basilieuterus rivularis"
/organella="mitochondrion"
/mol_type="genomic DNA"
/specimen_voucher="Louisiana State University Museum of
Natural Science avian tissue B1146"
/db_xref="taxon:240589"
1..1041
/note="ND2"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 2"
/protein_id="AA32825.1"
/db_xref="GI:39979798"
/translation="NM004511FITSLSGTTTSSNMWMTGLINLAIPLI
SCKHRAIEAKYFVLTQAAALVLFSSNMWMTGLINLAIPLI
AKGLVPHFPPPEVLOSPLTGLISTIKLSPITLLVWTSYSLNPSSILALAI
STLGGWGLNTOIKKILAFSSISHLGMAITTYNPKITLINYIVAMATVET
LNTKVLKSLTMTATKIPSLNAMLLTLSDAGLPVLTGLPKMLIOELTKDVA
PATLISLSLSLTFYRLAYCTITLPPHTNNKQWRANKSTNMAIILITNSVM
LIPISMLITIV"
1400..2950
/note="COI"

RESULT 4
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
SOURCE
ORIGIN
Query Match
Best Local Similarity 92.4%; Score 19.4; DB 5; Length 4692;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CCAGTCAACAGTGTGT 21
1555 CCAGTCAACAGTGTGT 1575

RESULT 4
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
SOURCE
ORIGIN
Query Match
Best Local Similarity 95.2%; Score 19.4; DB 5; Length 4692;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

AY327393 4692 bp DNA linear VRT 23-JUL-2004
Basilieuterus fulvicauda NADH dehydrogenase subunit 2, cytochrome
oxidase subunit I, cytochrome oxidase subunit II, ATP-synthase 8,
and ATP-synthase 6 genes, complete cds; mitochondrial.
AY327393.1 GI:39979809
mitochondrion Basilieuterus fulvicauda (buff-rumped warbler)
Basilieuterus fulvicauda
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
Fringillidae; Emberizinae; Basilieuterus.
1 (bases 1 to 4692)
Lovelte, J. J.
Molecular phylogeny and plumage signal evolution in a trans Andean
and circum Amazonian avian species complex
Mol. Phylogenet. Evol. 32 (2), 512-523 (2004)
2 (bases 1 to 4692)
Lovelte, J. J.
Direct Submission
Submitted (20-JUN-2003) Laboratory of Ornithology, Cornell
University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA
Location/Qualifiers
1..4692
/organism="Basilieuterus fulvicauda"
/organella="mitochondrion"
/mol_type="genomic DNA"
/specimen_voucher="Louisiana State University Museum of
Natural Science avian tissue B1146"
/db_xref="taxon:240589"
1..1041
/note="ND2"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 2"
/protein_id="AA32825.1"
/db_xref="GI:39979798"
/translation="NM004511FITSLSGTTTSSNMWMTGLINLAIPLI
SCKHRAIEAKYFVLTQAAALVLFSSNMWMTGLINLAIPLI
AKGLVPHFPPPEVLOSPLTGLISTIKLSPITLLVWTSYSLNPSSILALAI
STLGGWGLNTOIKKILAFSSISHLGMAITTYNPKITLINYIVAMATVET
LNTKVLKSLTMTATKIPSLNAMLLTLSDAGLPVLTGLPKMLIOELTKDVA
PATLISLSLSLTFYRLAYCTITLPPHTNNKQWRANKSTNMAIILITNSVM
LIPISMLITIV"
1400..2950
/note="COI"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:50:19 ; Search time 34.5345 Seconds
(without alignments)
3127.876 Million cell updates/sec

Title: US-10-088-666-74

Perfect score: 20

Sequence: 1 gaggaagaagttctctat 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3625171 seqs, 2700493622 residues 7250342

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	92.0	6109	10	US-09-795-061-1
2	18	90.0	465	18	US-10-425-115-45008
3	17.4	87.0	593	15	US-10-029-386-1346
4	17.4	87.0	6200	10	US-09-795-061-3
5	17	85.0	127267	17	US-10-322-281-797
6	16.8	84.0	175561	14	US-10-017-721-3
7	16.8	84.0	175561	16	US-10-235-192A-48
8	16.8	84.0	1230025	16	US-10-289-762-1
9	16.8	84.0	1503841	9	US-09-795-668-1
10	16.8	84.0	1503841	9	US-09-795-686-1
11	16.8	84.0	1503841	9	US-09-346-807-1
12	16.4	82.0	255	15	US-10-149-736-7

C 13	16.4	82.0	465	9	US-09-974-300-1058	Sequence 3058, Ap
C 14	16.4	82.0	688	13	US-10-027-632-102170	Sequence 102170,
C 15	16.4	82.0	688	13	US-10-027-632-102171	Sequence 102171,
C 16	16.4	82.0	688	15	US-10-027-632-102170	Sequence 102170,
C 17	16.4	82.0	688	15	US-10-027-632-102171	Sequence 102171,
C 18	16.4	82.0	1340	10	US-09-845-416-11	Sequence 11, Appl
C 19	16.4	82.0	1667	10	US-09-845-416-7	Sequence 7, Appl
C 20	16.4	82.0	1991	10	US-09-845-416-3	Sequence 3, Appl
C 21	16.4	82.0	2022	17	US-10-437-963-57769	Sequence 57769, A
C 22	16.4	82.0	3446	10	US-09-845-416-14	Sequence 14, Appl
C 23	16.4	82.0	3510	10	US-09-845-416-12	Sequence 12, Appl
C 24	16.4	82.0	3531	10	US-09-845-416-10	Sequence 10, Appl
C 25	16.4	82.0	3858	10	US-09-845-416-9	Sequence 9, Appl
C 26	16.4	82.0	3999	10	US-09-845-416-6	Sequence 6, Appl
C 27	16.4	82.0	4182	10	US-09-845-416-2	Sequence 2, Appl
C 28	16.4	82.0	4414	10	US-09-845-416-32	Sequence 32, Appl
C 29	16.4	82.0	4476	10	US-09-845-416-31	Sequence 31, Appl
C 30	16.4	82.0	4498	10	US-09-845-416-30	Sequence 30, Appl
C 31	16.4	82.0	4825	10	US-09-845-416-29	Sequence 29, Appl
C 32	16.4	82.0	4848	10	US-09-845-416-35	Sequence 35, Appl
C 33	16.4	82.0	4966	10	US-09-845-416-28	Sequence 28, Appl
C 34	16.4	82.0	4990	10	US-09-845-416-34	Sequence 34, Appl
C 35	16.4	82.0	5060	10	US-09-845-416-36	Sequence 36, Appl
C 36	16.4	82.0	5149	10	US-09-845-416-27	Sequence 27, Appl
C 37	16.4	82.0	5339	15	US-10-149-736-40	Sequence 40, Appl
C 38	16.4	82.0	5417	15	US-10-149-736-39	Sequence 39, Appl
C 39	16.4	82.0	5462	15	US-10-149-736-41	Sequence 41, Appl
C 40	16.4	82.0	8689	15	US-10-149-736-42	Sequence 42, Appl
C 41	16.4	82.0	11058	10	US-09-845-416-1	Sequence 1, Appl
C 42	16.4	82.0	11443	15	US-10-149-736-44	Sequence 44, Appl
C 43	16.4	82.0	12057	15	US-10-149-736-47	Sequence 47, Appl
C 44	16.4	82.0	13957	9	US-09-782-378A-22	Sequence 22, Appl
C 45	16.4	82.0	13957	9	US-09-880-107-2284	Sequence 2284, Ap

ALIGNMENTS

RESULT 1
US-09-795-061-1, Application US/09795061
; Sequence 1, Imamura, Yasutada
; Publication No. US20030166842A1
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Imamura, Yasutada
; TITLE OF INVENTION: Pro-Alpha 3 (V) Collagen Genes
; FILE REFERENCE: 960296.96781
; CURRENT APPLICATION NUMBER: US/09/795, 061
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6109
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (82)..(5298)
US-09-795-061-1

Query Match 92.0%; Score 18.4; DB 10; Length 6109;
Best Local Similarity 95.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 1; Indels 0;
Gaps 0;
Gy 1 GAGGGAAGAAGTCTCTAT 20
Db 4907 GAGGGAAGAAGTCTCTAT 4926
RESULT 2
US-10-425-115-45008
; Sequence 45008, Application US/10425115
; Publication No. US20040214272A1

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovall, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 45008
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MNT4577_141051C.1
US-10-425-115-45008

```

```

Query Match      90.0%; Score 18; DB 18; Length 465;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 GGGAGAGAGTCTCTTAT 20
DB      439 GGGAGAGAGTCTCTTAT 456

```

```

RESULT 3
US-10-029-386-1346
; Sequence 1346, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hatzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: A60MCA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1346
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC013816.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.56
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.51
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.43
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.46
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.36
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; OTHER INFORMATION: SWISSPROT HIT: Q00083, EVALU8 5.80e+00
; OTHER INFORMATION: EST HUMAN HIT: S6539472.1, EVALU8 8.30e-01
; OTHER INFORMATION: NT HIT: AJ314911.1, EVALU8 2.20e+00
US-10-029-386-1346

```

```

Query Match      87.0%; Score 17.4; DB 15; Length 593;
Best Local Similarity 94.7%; Pred. No. 74;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 AGGAGAGAGTCTCTTAT 20
DB      524 AGGAGAGAGTCTCTTAT 542

```

```

RESULT 4
US-09-795-061-3
; Sequence 3, Application US/09795061
; Publication No. US20030166842A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Imamura, Yasutada
; TITLE OF INVENTION: Pro-Alpha 3 (V) Collagen Genes
; FILE REFERENCE: 960296.96781
; CURRENT APPLICATION NUMBER: US/09/795,061
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(5321)
US-09-795-061-3

```

```

Query Match      87.0%; Score 17.4; DB 10; Length 6200;
Best Local Similarity 94.7%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GAGGAGAGAGTCTCTTAT 19
DB      4927 GAGGAGAGAGTCTCTTAT 4945

```

```

RESULT 5
US-10-322-281-797
; Sequence 797, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945200100
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 797
; LENGTH: 127767
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(127767)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-797

```

```

Query Match      85.0%; Score 17; DB 17; Length 127767;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 GAGAGAGTCTCTTAT 20
DB      5129 GAGAGAGTCTCTTAT 5145

```

```

RESULT 6
US-10-017-721-3/C
; Sequence 3, Application US/10017721
; Publication No. US20030096248A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Daley, George
; APPLICANT: Bolx, Stacey
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MMI-003
; CURRENT APPLICATION NUMBER: US/10/017,721
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/317,033
; PRIOR FILING DATE: 2001-09-04

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:14 ; Search time 8.64865 Seconds
(without alignments)
1643.659 Million cell updates/sec

Title: US-10-088-666-74

Perfect score: 20

Sequence: 1 gaggaagaagtcctctat 20

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCITUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	6109	4	US-09-795-061-1 Sequence 1, Appli
2	17.4	87.0	6200	4	US-09-795-061-3 Sequence 3, Appli
3	16.8	84.0	1230025	4	US-09-198-452A-1 Sequence 1, Appli
4	16.4	82.0	5952	4	US-09-687-875A-1 Sequence 1, Appli
5	15.4	82.0	13977	4	US-09-484-970B-60 Sequence 60, Appli
6	15.8	79.0	239	3	US-09-443-184-18 Sequence 18, Appli
7	15.8	79.0	244	3	US-09-443-184-19 Sequence 19, Appli
8	15.8	79.0	267	3	US-09-443-184-17 Sequence 17, Appli
9	15.8	79.0	3001	4	US-09-539-333D-156 Sequence 156, App
10	15.8	79.0	3182	1	US-07-797-556-5 Sequence 5, Appli
11	15.8	79.0	3182	1	US-07-943-843-1 Sequence 1, Appli
12	15.8	79.0	3182	1	US-08-347-003-1 Sequence 1, Appli
13	15.8	79.0	3591	1	US-07-943-843-5 Sequence 5, Appli
14	15.8	79.0	3591	1	US-08-347-003-5 Sequence 5, Appli
15	15.8	79.0	26000	4	US-09-843-376-10 Sequence 10, Appli
16	15.4	77.0	4519	4	US-08-976-259-6 Sequence 6, Appli
17	15.4	77.0	4519	4	US-09-956-004-6 Sequence 6, Appli
18	15.4	77.0	4712	4	US-09-221-017B-577 Sequence 577, App
19	15.4	77.0	148567	4	US-09-801-876B-3 Sequence 3, Appli
20	15.4	77.0	148567	4	US-10-254-869-3 Sequence 3, Appli
21	15.2	76.0	2344	4	US-08-893-852A-2 Sequence 2, Appli
22	15.2	76.0	2432	1	US-08-339-214-29 Sequence 29, Appli
23	15.2	76.0	2808	2	US-08-153-721B-1 Sequence 1, Appli
24	15.2	76.0	2808	2	US-09-427-154-1 Sequence 1, Appli
25	15.2	76.0	3394	4	US-09-350-982C-4 Sequence 4, Appli
26	15.2	76.0	3498	4	US-09-350-982C-4 Sequence 4, Appli
27	15.2	76.0	3515	4	US-09-221-017B-824 Sequence 824, App

ALIGNMENTS

28	15.2	76.0	3797	4	US-09-696-668-1 Sequence 1, Appli
29	15.2	76.0	3816	4	US-09-696-668-2 Sequence 2, Appli
30	15.2	76.0	4275	4	US-09-972-115A-5 Sequence 5, Appli
31	15.2	76.0	4512	4	US-09-350-982C-3 Sequence 3, Appli
32	15.2	76.0	5075	4	US-09-799-451-556 Sequence 556, App
33	15.2	76.0	6018	4	US-09-849-602-111 Sequence 11, Appli
34	15.2	76.0	6114	4	US-09-543-681A-1366 Sequence 1366, Ap
35	15.2	76.0	128779	4	US-09-497-855A-38 Sequence 38, Appli
36	15	75.0	1166	4	US-09-325-932A-139 Sequence 139, App
37	14.8	74.0	597	2	US-08-332-766A-19 Sequence 19, Appli
38	14.8	74.0	990	4	US-09-134-000C-1245 Sequence 1245, Ap
39	14.8	74.0	1025	3	US-08-643-212-17 Sequence 17, Appli
40	14.8	74.0	1025	3	US-08-643-212-19 Sequence 19, Appli
41	14.8	74.0	1025	3	US-08-643-212-21 Sequence 21, Appli
42	14.8	74.0	1025	3	US-08-643-212-23 Sequence 23, Appli
43	14.8	74.0	1025	3	US-08-643-212-25 Sequence 25, Appli
44	14.8	74.0	1025	3	US-08-643-212-27 Sequence 27, Appli
45	14.8	74.0	1025	3	US-08-643-212-29 Sequence 29, Appli

RESULT 1
US-09-795-061-1
Sequence 1, Application US/09795061
Patent No. 6759528
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
TITLE OF INVENTION: Pro-Alpha 3 (V) Collagen Genes
FILE REFERENCE: 960296, 96781
CURRENT FILING DATE: 2001-02-26
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 6109
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (82)..(5298)
US-09-795-061-1

Query Match 92.0%; Score 18.4; DB 4; Length 6109;
Best Local Similarity 95.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGAAGAAGTCTCTTAT 20
Db 4907 GAGGAAGAAGTCTCTTAT 4926

RESULT 2
US-09-795-061-3
Sequence 3, Application US/09795061
Patent No. 6759528
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
TITLE OF INVENTION: Pro-Alpha 3 (V) Collagen Genes
FILE REFERENCE: 960296, 96781
CURRENT FILING DATE: 2001-02-26
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 6200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS

```
LOCATION: (87) .. (5321)
US-09-795-061-3
Query Match 87.0%; Score 17.4; DB 4; Length 6200;
Best Local Similarity 94.7%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGAGAAAGTCTCTTA 19
Db 4927 GAGGAGAAAGTCTCTTA 4945

RESULT 3
US-09-198-452A-1/c
Sequence 1, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
NAME/KEY: misc_feature
LOCATION: (1) .. (15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (15001) .. (30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (30001) .. (45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (45001) .. (60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (60001) .. (75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (75001) .. (90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (90001) .. (105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (105001) .. (120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (120001) .. (135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (135001) .. (150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (150001) .. (165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (165001) .. (180000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (180001) .. (195000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (195001) .. (210000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

LOCATION: (210001) .. (225000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (225001) .. (240000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (240001) .. (255000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (255001) .. (270000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (270001) .. (285000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (285001) .. (300000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (300001) .. (315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (315001) .. (330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (330001) .. (345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001) .. (360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001) .. (375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001) .. (390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001) .. (405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001) .. (420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001) .. (435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001) .. (450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001) .. (465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001) .. (480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001) .. (495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001) .. (510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001) .. (525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001) .. (540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001) .. (555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001) .. (570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001) .. (585000)
```


CC microbes, either as a general test for contamination or as a test
CC specific for particular genera or (sub)species. It is quicker than known
CC microbiological methods, and can detect several organisms in the same
CC sample, including organisms not presently recognized as contaminants. The
CC method provides an early indication of contamination and can be automated
CC for high throughput analysis

XX Sequence 20 BP; 6 A; 2 C; 6 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGAGAGAGTCTCTAT 20
Db 1 GAGGAGAGAGTCTCTAT 20

RESULT 2
AA61566
ID AA61566 standard; DNA; 267 BP.

XX AA61566;

XX 02-JUL-2001 (first entry)

XX Lactobacillus brevis 23S rRNA-spacer-5S rRNA DNA fragment.

XX 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

XX Lactobacillus brevis.

XX DB19945964-A1.

XX 05-APR-2001.

XX 24-SEP-1999; 99DE-01045964.

XX 24-SEP-1999; 99DE-01045964.

XX (BIOT-) BIOTECON DIAGNOSTICS GMBH.

XX Fandke M, Gasch A, Berghof K;

XX WPI; 2001-246136/26.

XX Detecting contaminating microorganisms in brewing, by nucleic acid
PT amplification and hybridization, either non-specific or genus- or species
PT -specific.

XX Claim 9(1); Page 9; 48pp; German.

XX This invention describes a novel method for detecting microorganisms (A)
CC of importance in brewing which comprises treating a sample with at least
CC two primers (P1) that hybridize to a consensus region in the nucleic acid
CC of (A), at least part of the microbial nucleic acid is amplified, the
CC amplicon is treated with at least one probe (P2) that hybridizes
CC specifically with a sequence common to all (A) or specific for one or
CC more families, genera or species, and any formation of hybrids is
CC detected. The method is used to detect, identify and/or characterize
CC microorganisms in beer or brewing materials, particularly for detecting
CC contamination. The method may detect the entire range of contaminating
CC microbes, either as a general test for contamination or as a test
CC specific for particular genera or (sub)species. It is quicker than known
CC microbiological methods, and can detect several organisms in the same
CC sample, including organisms not presently recognized as contaminants. The
CC method provides an early indication of contamination and can be automated
CC for high throughput analysis

XX Sequence 267 BP; 77 A; 51 C; 77 G; 62 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 5; Length 267;
Best Local Similarity 100.0%; Pred. No. 6.1;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAGGAGAGAGTCTCTAT 20
Db 154 GAGGAGAGAGTCTCTAT 173

RESULT 3
ABN26443
ID ABN26443 standard; cDNA; 337 BP.

XX ABN26443;

XX 24-JUN-2002 (first entry)

XX Human ORFX polynucleotide sequence SEQ ID NO:21363.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis; gene; ss.

XX Homo sapiens.

XX W0200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US010836.

XX 30-MAY-2000; 2000US-0206132P.

XX 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shinketsu RA, Leach MD;

XX WPI; 2002-106308/14.

XX P-PSDB; ABP106591.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.

XX Disclosure; SEQ ID NO 21363; 1037pp; English.

XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification)). ABN15762 to ABN27253 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic

RESULT 1						PAT 30-APR-2001
LOCUS AXI05802	20 bp	DNA				
DEFINITION Sequence 74 from Patent WO0123605.						
ACCESSION AXI05802						
VERSION AXI05802.1 GI:13921815						
KEYWORDS .						
SOURCE synthetic construct ORGANISM synthetic construct artificial sequences.						
REFERENCE 1						
AUTHORS Fandke,M., Gasch,A. and Berghof,K. TITLE Method and nucleic acids for determining the presence of micro-organisms specific to the brewing process JOURNAL Patent: WO 0123605-A 74 05-APR-2001; Biotecon Diagnostische GmbH (DE) FEATURES Location/Qualifiers source 1..20 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="spezifische Sequenz fur Lactobacillus brevis"						
ORIGIN						
Query Match	100.0%;	Score 20; DB 6;	Length 20;			
Best Local Similarity	100.0%;	Fed. No. 1.3e+02;				
Matches 20; Conservative 0;	Mismatches 0;	Gaps 0;				
OY	1 GAGGGAAGAAGTCTCCTTAAT 20 1 GAGGGAAGAAGTCTCCTTAAT 20					
Db	1 GAGGGAAGAAGTCTCCTTAAT 20					
RESULT 2						
LOCUS AXI05729	267 bp	DNA				
DEFINITION Sequence 1 from Patent WO0123605.						
ACCESSION AXI05729						
VERSION AXI05729.1 GI:13921742						
KEYWORDS .						
SOURCE lactobacillus brevis ORGANISM lactobacillus brevis						

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 ; Search time 417.042 Seconds
(without alignments)
2184.416 Million cell updates/sec

Title: US-10-088-666-73
Perfect score: 25
Sequence: 1 tcgaataatcgaataatcctag 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 1821985908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gse1:
9: gb_gse2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.8	83.2	717	9	AC296763 Mus muscu
2	20.2	80.8	460	9	AQ489622 RPT-11-2
3	19.8	79.2	887	9	CNS06M55
4	19.4	77.6	848	8	AZ683885 ENTP32TF
5	19.4	77.6	877	8	CC141515 NDL.32X2
6	19.4	77.6	900	8	BH133066 ENTPC66TR
7	19.4	77.6	900	8	BH147215 ENTQP4ATR
8	19.4	77.6	903	8	BH136872 ENTNA82TR
9	19.2	76.8	214	9	CR398469 Arabidops
10	19.2	76.8	277	9	CG148828 PUFKVA6TB
11	19.2	76.8	307	9	CL257264 ZMMBD061
12	19.2	76.8	340	6	CD324627 Str-Pus37
13	19.2	76.8	382	9	CE401158 tigr-gss-
14	19.2	76.8	385	9	CE624663 tigr-gss-
15	19.2	76.8	390	8	BZ632366 PUFKVA4TD
16	19.2	76.8	405	9	CG148831 PUFKVA4TD
17	19.2	76.8	409	8	BZ632406 PUFKVA4TD
18	19.2	76.8	413	8	BZ632362 PUFKVA4TD
19	19.2	76.8	424	8	BZ632408 PUFKVA4TD
20	19.2	76.8	464	9	BH170927 SALX_0035
21	19.2	76.8	464	9	CR853568 tigr-gss-
22	19.2	76.8	482	7	CG517136 s13dSG29H
23	19.2	76.8	568	9	CE345622 tigr-gss-
24	19.2	76.8	570	9	CE313026 tigr-gss-

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
25	19.2	76.8	583	9	CE029479			Hattori, M., Toyoda, A., Noguichi, H., Kojima, T. and Sakaki, Y.	BAC end Sequences of Library MSWg01	
26	19.2	76.8	601	9	CE378676			Unpublished		
27	19.2	76.8	630	9	CE407289			2 (bases 1 to 717)		
28	19.2	76.8	639	9	CE397029			Hattori, M., Toyoda, A., Noguichi, H., Kojima, T. and Sakaki, Y.	Direct Submission	
29	19.2	76.8	646	9	CE717195			Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan		
30	19.2	76.8	685	9	CE745651			tel:81-45-503-9111, Fax:81-45-503-9170		
31	19.2	76.8	709	9	CE827541			lib:81-45-503-9111, Fax:81-45-503-9170		
32	19.2	76.8	711	9	AG123355			Clones are derived from the mouse BAC library MSWg01. For BAC library availability, please contact Kunhya Abe (abe@ric.riken.jp).		
33	19.2	76.8	742	9	CE459772			The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyada, Tsukuba, 305-0074 Japan		
34	19.2	76.8	743	9	CE224657			phone: 81-298-36-9185, Fax: 81-298-36-9199		
35	19.2	76.8	752	9	CG048515			e-mail: abe@ric.riken.jp		
36	19.2	76.8	833	9	CG048515			PRIMERS		
37	19.2	76.8	888	9	BZ829021			Sequencing : T7		
38	19.2	76.8	888	9	CG965204			LIBRARY		
39	18.8	75.2	402	8	AQ207867			Vector 1 : BAC3.6		
40	18.8	75.2	405	1	A1872073			R Site 1 : BCR1		
41	18.8	75.2	506	9	CG696733			R Site 2 : BCR1		
42	18.8	75.2	574	8	AQ466749			Location/Qualifiers		
43	18.8	75.2	595	9	CE163486			1. 717		
44	18.8	75.2	628	6	CD004668			/organism="Mus musculus molossinus"		
45	18.8	75.2	630	9	CR323914			/mol_type="genomic DNA"		
								/sub_species="molossinus"		
								/db_xref="taxon:57486"		
								/clone="MSWg01-073H13.T7"		
								/sex="male"		
								/tissue_type="mixture of kidney and spleen"		
								/clone_lib="MSWg01 Mouse Male BAC Library"		

FEATURES

source

ORIGIN
 Query Match 83.2%; Score 20.8; DB 9; Length 717;
 Best Local Similarity 91.7%; Pred. No. 5.3e+02;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCTA 24
 |||||
 DB 507 TCGAGATATATGATATATCTA 484

RESULT 2
 A0489622 460 bp DNA linear GSS 24-APR-1999
 LOCUS RPCI-11-231C16.TV RPCI-11 Homo sapiens genomic clone
 DEFINITION RPCI-11-231C16, genomic survey sequence.
 ACCESSION A0489622
 VERSION A0489622.1 GI:4675496
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 460)
 Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and
 Venter, J.C.
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 Unpublished (1997)
 Other_GSS: RPCI-11-231C16.TV
 Contact: Shaying Zhao, William Niernan, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeet@ig.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genet cs (info@resgen.com). BAC end search page:
 http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: T7
 Class: BAC ends.
 FEATURES
 source
 Location/Qualifiers
 1..460
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7588383"
 /db_xref="taxon:9606"
 /clone_lib="RPCI-11-231C16"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: PBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
 RPCI11 Human Male BAC Library"

ORIGIN
 Query Match 80.8%; Score 20.2; DB 8; Length 460;
 Best Local Similarity 88.0%; Pred. No. 9.2e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCTAG 25
 |||||
 DB 250 TCGAGAGAGCTTGATATATCTAG 226

RESULT 3
 CNS06W55 887 bp DNA linear GSS 06-JUL-2001
 LOCUS T3 end of clone AX0AA028C03 of library AX0AA from strain CBS 7064
 DEFINITION of Pichia farinosa, genomic survey sequence.

ACCESSION AL417999
 VERSION AL417999.1 GI:12200400
 KEYWORDS GSS.
 SOURCE Pichia farinosa
 ORGANISM Pichia farinosa
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Pichia.
 1 (bases 1 to 887)
 Souciet, J., Algle, M., Artiguenave, F., Blandin, G.,
 Souciet, J., L., Algle, M., Artiguenave, F., Blandin, G.,
 Bolotin-Fukuhara, M., Bon, E., Brotier, P., Casaregola, S.,
 de Montigny, J., Dujon, B., Durand, P., Lepingle, A., Lorente, B.,
 Malpertuy, A., Neuvéglise, C., Olier-Kalogeropoulos, O., Potier, S.,
 Sautin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Jouvel, M.,
 Wincker, P. and Weissenbach, J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 FEMS Lett. 487 (1), 3-12 (2000)

TITLE
 JOURNAL MEDLINE
 PUBMED 11152876
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL MEDLINE
 PUBMED 20584725
 2 (bases 1 to 887)
 de Montigny, J., Spehner, C., Souciet, J., Tekala, F., Dujon, B.,
 Wincker, P., Artiguenave, F. and Potier, S.
 Genomic exploration of the hemiascomycetous yeasts: 15. Pichia
 sorbitophila
 FEMS Lett. 487 (1), 87-90 (2000)

COMMENT
 Direct Submission
 Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazii, Zygosaccharomyces rouxi,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source
 Location/Qualifiers
 1..887
 /organism="Pichia farinosa"
 /mol_type="genomic DNA"
 /strain="CBS 7064"
 /db_xref="taxon:4920"
 /clone_lib="AX0AA028C03"
 /clone_lib="AX0AA"
 /note="end : T3"

ORIGIN
 Query Match 79.2%; Score 19.8; DB 9; Length 887;
 Best Local Similarity 84.0%; Pred. No. 1.3e+03;
 Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCTAG 25
 |||||
 DB 850 TCGAGATATATGATATATCTAG 826

RESULT 4
 A2683885 848 bp DNA linear GSS 14-DEC-2000
 LOCUS ENTLP32TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION genomic, genomic survey sequence.
 ACCESSION A2683885
 VERSION A2683885.1 GI:11821031
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica
 ORGANISM Entamoeba histolytica

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:50:19 ; Search time 43.1682 Seconds
(without alignments)
3127.876 Million cell updates/sec

Title: US-10-088-666-73

Perfect score: 1 tcgagatattgatactatctag 25

Sequence: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Scoring table: 3625171 seqs, 2700493622 residues

Searched: Total number of hits satisfying chosen parameters: 7250342

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US11_NEM_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18.8	75.2	228 14	US-10-016-349A-108
2	18.8	75.2	1324 14	US-10-016-349A-109
3	18.6	74.4	1206 18	US-10-425-115-119404
4	18.4	73.6	1891 17	US-10-437-963-45894
5	18.2	72.8	332 15	US-10-369-493-29816
6	18.2	72.8	588 15	US-10-029-386-6066
7	18.2	72.8	638 17	US-10-767-701-26220
8	18.2	72.8	2179 9	US-09-070-927A-468
9	18.2	72.8	163701 17	US-10-322-281-439
10	18.2	72.8	1830121 14	US-10-329-960-1
11	18.2	72.8	1830121 16	US-10-329-670-1
12	18.2	72.8	1830121 18	US-10-158-865-1

Query Match	Score	DB	Length	Best Local Similarity	Pred. No.	Indels	Gaps
Matches 20; Conservative	75.2%	18.8	228	90.9%	0	0	0
4 AGAATATGATATATCTAG 25							
17 AGAATGATGATATCTAG 38							

RESULT 2
US-10-016-349A-109/c

ALIGNMENTS

US-10-016-349A-108
Sequence 108, Application US/10016349A
Publication No. US2003006437B1
GENERAL INFORMATION:
APPLICANT: Recipon, Herve
APPLICANT: Sun, Yongming
APPLICANT: Chen, Sei-Yu
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods relating to Lung Specific Genes and Proteins
FILE REFERENCE: DEX-0243
CURRENT APPLICATION NUMBER: US/10/016,349A
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 60/243,459
PRIOR FILING DATE: 2000-10-26
NUMBER OF SEQ ID NOS: 244
SOFTWARE: PatentIn version 3.1
SEQ ID NO 108
LENGTH: 228
TYPE: DNA
ORGANISM: Homo sapiens
US-10-016-349A-108

```
/ Sequence 109, Application US/10016349A
/ Publication No. US20030064378A1
/ GENERAL INFORMATION:
/ APPLICANT: Recipon, Herve
/ APPLICANT: Sun, Yongming
/ APPLICANT: Chen, Sei-yu
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Turner, Leah
/ TITLE OF INVENTION: Compositions and Methods relating to Lung Specific Genes and Prot
/ FILE REFERENCE: DSX-0243
/ CURRENT APPLICATION NUMBER: US/10/016,349A
/ CURRENT FILING DATE: 2001-10-26
/ PRIOR APPLICATION NUMBER: US 60/243,459
/ PRIOR FILING DATE: 2000-10-26
/ NUMBER OF SEQ ID NOS: 244
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 109
/ LENGTH: 1324
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312)..(312)
/ OTHER INFORMATION: n = a, c, g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (385)..(385)
/ OTHER INFORMATION: n = a, c, g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (419)..(419)
/ OTHER INFORMATION: n = a, c, g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (506)..(506)
/ OTHER INFORMATION: n = a, c, g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (517)..(518)
/ OTHER INFORMATION: n = a, c, g or t
/ US-10-016-349A-109
/ Query Match 75.2%; Score 18.8; DB 14; Length 1324;
/ Best Local Similarity 90.9%; Pred. No. 1.2e+03;
/ Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AGAATATTGAATATATCTG 25
DB 1224 AGAATGATTGATGATCTG 1203

RESULT 3
/ US-10-425-115-119404/C
/ Sequence 119404, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 119404
/ LENGTH: 1206
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: WRT4577_40380C.1
/ US-10-425-115-119404
```

```
Query Match 74.4%; Score 18.6; DB 18; Length 1206;
Best Local Similarity 84.0%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCGGATATTGAATATATCTG 25
DB 271 TCGGATATTGAATATATCTG 247

RESULT 4
/ US-10-437-963-45894/C
/ Sequence 45894, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 45894
/ LENGTH: 1891
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_WRT4530_48814C.1
/ US-10-437-963-45894
/ Query Match 73.6%; Score 18.4; DB 17; Length 1891;
/ Best Local Similarity 95.0%; Pred. No. 1.8e+03;
/ Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AGAATATTGAATATATCT 23
DB 1795 AGAATATTGAATATATCT 1776

RESULT 5
/ US-10-369-493-29816/C
/ Sequence 29816, Application US/10369493
/ Publication No. US2003023675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 29816
/ LENGTH: 332
/ TYPE: DNA
/ ORGANISM: Caenorhabditis elegans
/ US-10-369-493-29816
/ Query Match 72.8%; Score 18.2; DB 15; Length 332;
/ Best Local Similarity 87.0%; Pred. No. 1.8e+03;
/ Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:14 ; Search time 10.8108 Seconds
(without alignments)
1643.699 Million cell updates/sec

Title: US-10-088-666-73

Perfect score: 25
Sequence: 1 tcgagataatgataatcag 25

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents NA: *
2: /cgm2_6/ptodata/1/ina/5A.COMB.seq: *
3: /cgm2_6/ptodata/1/ina/5B.COMB.seq: *
4: /cgm2_6/ptodata/1/ina/6A.COMB.seq: *
5: /cgm2_6/ptodata/1/ina/6B.COMB.seq: *
6: /cgm2_6/ptodata/1/ina/6C.COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	73.6	5064	US-08-936-107A-8	Sequence 8, Appl1
C 2	18.2	72.8	234	US-09-134-000C-317	Sequence 317, App
C 3	18.2	72.8	468	US-09-134-000C-316	Sequence 316, App
C 4	18.2	72.8	1830121	US-09-557-884-1	Sequence 1, Appl1
C 5	18.2	72.8	1830121	US-09-643-990A-1	Sequence 1, Appl1
C 6	18.2	72.8	1830121	US-10-329-960-1	Sequence 1, Appl1
C 7	17.6	70.4	103	US-09-313-284A-1738	Sequence 1738, Ap
C 8	17.6	70.4	1001	US-09-313-294A-2255	Sequence 2255, Ap
C 9	17.6	70.4	1001	US-09-641-638-111	Sequence 111, App
C 10	17.6	70.4	1001	US-09-641-638-112	Sequence 112, App
C 11	17.6	70.4	1001	US-10-170-097-111	Sequence 111, App
C 12	17.6	70.4	1001	US-10-170-097-112	Sequence 112, App
C 13	17.4	69.6	1830121	US-09-557-884-1	Sequence 1, Appl1
C 14	17.4	69.6	1830121	US-09-643-990A-1	Sequence 1, Appl1
C 15	17.4	69.6	1830121	US-10-329-960-1	Sequence 1, Appl1
C 16	17.2	68.8	441	US-09-710-279-1795	Sequence 1795, Ap
C 17	17.2	68.8	462	US-09-134-001C-2294	Sequence 2294, Ap
C 18	17.2	68.8	2943	US-09-710-279-3983	Sequence 3983, Ap
C 19	17.2	68.8	3249	US-09-710-279-3622	Sequence 3622, Ap
C 20	17.2	68.8	3442	US-09-710-279-4030	Sequence 4030, Ap
C 21	17.2	68.8	1557	US-09-270-767-28519	Sequence 28519, A
C 22	17.2	68.0	1585	US-09-270-767-28996	Sequence 28996, A
C 23	17.2	68.0	1907	US-09-270-767-13100	Sequence 13100, A
C 24	17.2	68.0	2970	US-09-710-279-2593	Sequence 2593, Ap
C 25	17.2	68.0	3128	US-09-270-767-12706	Sequence 12706, Ap
C 26	17.2	68.0	3271	US-09-710-279-4142	Sequence 4142, Ap
C 27	17.2	68.0	3461	US-09-710-279-3713	Sequence 3713, Ap

C 28	16.8	67.2	222	US-09-328-352-1425	Sequence 1425, Ap
C 29	16.8	67.2	564	US-09-248-796A-2085	Sequence 2085, Ap
C 30	16.8	67.2	759	US-09-710-279-1031	Sequence 1031, Ap
C 31	16.8	67.2	834	US-09-248-796A-2217	Sequence 2217, Ap
C 32	16.8	67.2	1386	US-09-134-001C-2474	Sequence 2474, Ap
C 33	16.8	67.2	2301	US-09-710-279-419	Sequence 419, App
C 34	16.8	67.2	2634	US-09-710-279-4069	Sequence 4069, App
C 35	16.8	67.2	2744	US-09-710-279-3892	Sequence 3892, Ap
C 36	16.8	67.2	3170	US-09-710-279-4058	Sequence 4058, Ap
C 37	16.8	67.2	3321	US-09-107-532A-3826	Sequence 3028, Ap
C 38	16.8	67.2	3552	US-09-710-279-3430	Sequence 3430, Ap
C 39	16.8	67.2	3825	US-09-710-279-3799	Sequence 3799, App
C 40	16.8	67.2	3968	US-09-710-279-759	Sequence 759, App
C 41	16.8	67.2	7215	US-09-134-001C-627	Sequence 627, App
C 42	16.8	67.2	7215	US-09-513-999C-10571	Sequence 10571, A
C 43	16.6	66.4	735	US-09-107-532A-1648	Sequence 1648, Ap
C 44	16.6	66.4	735	US-08-956-171B-178	Sequence 178, App
C 45	16.6	66.4	3025	US-08-956-171B-178	Sequence 178, App

ALIGNMENTS

RESULT 1
US-08-936-107A-8/c
Sequence 8, Application US/08936107A
Patent No. 6403306
GENERAL INFORMATION:
APPLICANT: Stephens, David S.
TITLE OF INVENTION: Serogroup-Specific Nucleotide Sequences
TITLE OF INVENTION: In the Molecular Typing of Bacterial Isolates and the
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESS: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,107A
FILING DATE: 23-SEP-1997
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/827,622
FILING DATE: 09-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Caruthers, Jennie M.
REGISTRATION NUMBER: 34,464
REFERENCE/DOCKET NUMBER: 77-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8080
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5064 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULAR TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 479..1597
FEATURE:
NAME/KEY: CDS

LOCATION: 1599..3236
FEATURE:
NAME/KEY: CDS
LOCATION: 3309..4052
FEATURE:
NAME/KEY: CDS
LOCATION: 4054..4917
US-08-936-107A-8

Query Match 73.6%; Score 18.4; DB 4; Length 5064;
Best Local Similarity 95.0%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATA 20
DB 3612 TCGATATATATGATATATA 3593

RESULT 2
US-09-134-000C-317
Sequence 317, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patent version 3.1
SEQ ID NO 317
LENGTH: 234
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-317

Query Match 72.8%; Score 18.2; DB 4; Length 234;
Best Local Similarity 87.0%; Pred. No. 11e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCT 23
DB 191 TCGAGATATATTCATATATACT 213

RESULT 3
US-09-134-000C-316/C
Sequence 316, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patent version 3.1
SEQ ID NO 316
LENGTH: 468
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-316

Query Match 72.8%; Score 18.2; DB 4; Length 468;
Best Local Similarity 87.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCT 23
DB 211 TCGAGATATATTCATATATACT 189

RESULT 4
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and uses thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 72.8%; Score 18.2; DB 4; Length 1830121;
Best Local Similarity 87.0%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CGAGATATATGATATATCTA 24
DB 354497 CGAGATATATGATATCTTTA 354519

RESULT 5
US-09-643-990A-1
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and uses thereof
NUMBER OF SEQUENCES: 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 ; Search time 48.4985 Seconds
(without alignments)
2705.971 Million cell updates/sec

Title: US-10-088-666-73

Perfect score: 25
Sequence: 1 tcgagataatgataatcctag 25

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapept 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	25	100.0	25	AAF61638
2	25	100.0	267	AAF61566
3	18.8	75.2	228	ABO82965
4	18.8	75.2	1324	ABO82966
5	18.4	73.6	5064	AAV69110
6	18.4	73.6	25426	ADA41645
7	18.4	73.6	25426	ACC50940
8	18.4	73.6	25426	ADD38155
9	18.4	72.8	234	ADH82432
10	18.4	72.8	234	ADH82432
11	18.2	72.8	468	ADH82431
12	18.2	72.8	588	ACH72871
13	18.2	72.8	2179	AAH13405
14	18.2	72.8	2179	ABSS9200
15	18.2	72.8	22620	ABL07500
16	18.2	72.8	110000	AAH42063_03
17	18.2	72.0	1190	AAH1651
18	18.2	72.0	1190	ABSS9446
19	18.2	72.0	5962	ABL33286
20	17.8	71.2	1151	ADH82432
21	17.8	71.2	3393	ABH17845

22	17.6	70.4	103	10	ABX83278	ABx83278
23	17.6	70.4	160	6	ABL74381	AB174381
24	17.6	70.4	210	10	ABX83795	ABx83795
25	17.6	70.4	349	6	ABL83657	AB183657
26	17.6	70.4	1001	3	AAH57478	AAH57478
27	17.6	70.4	1001	3	AAH57477	AAH57477
28	17.6	70.4	1486	3	AAH39081	AAH39081
29	17.6	70.4	2076	8	AAH39067	AAH39067
30	17.6	70.4	2076	8	AAH39067	AAH39067
31	17.6	70.4	2345	2	AAH39061	AAH39061
32	17.6	70.4	3212	2	AAH39061	AAH39061
33	17.6	70.4	3212	10	AAH56488	AAH56488
34	17.6	70.4	5387	4	ABL21402	ABL21402
35	17.6	70.4	6032	6	AAH53311	AAH53311
36	17.6	70.4	8042	4	ABL22258	ABL22258
37	17.6	70.4	14429	6	ABL34242	ABL34242
38	17.6	70.4	14429	6	ABO67097	ABO67097
39	17.6	70.4	23989	4	AAH2555	AAH2555
40	17.6	70.4	110000	2	AAH20248_00	AAH20248
41	17.6	70.4	157050	12	ADH47194	ADH47194
42	17.6	70.4	312477	12	ADH69744	ADH69744
43	17.4	69.6	324	12	ADH35221	ADH35221
44	17.4	69.6	558	4	AAH53233	AAH53233
45	17.4	69.6	558	8	AAH34026	AAH34026

ALIGNMENTS

RESULT 1
ID AAF61638 standard; DNA: 25 BP.
AC AAF61638;
DT 02-JUL-2001 (first entry)
XX Lactobacillus brevis 23S rRNA/ss rRNA specific probe SRQ ID 73.
DE Lactobacillus brevis 23S rRNA/ss rRNA specific probe; brewing; beer; contamination; ss.
KM 23S rRNA, 5S rRNA; detection; probe; brewing; beer; contamination; ss.
XX Lactobacillus brevis.
OS Lactobacillus brevis.
XX DE19945964-A1.
XX 05-APR-2001.
XX 24-SEP-1999; 99DE-01045964.
XX 24-SEP-1999; 99DE-01045964.
XX 24-SEP-1999; 99DE-01045964.
XX (BIOT-) BIOTECON DIAGNOSTICS GMBH.
XX Fandke M, Gasch A, Bergdorf K;
XX WPI; 2001-246136/26.
XX Detecting contaminating microorganisms in brewing, by nucleic acid
XX amplification and hybridization, either non-specific or genus- or species
XX specific.
XX Claim 9(1); Page 17; 48pp; German.
XX
XX This invention describes a novel method for detecting microorganisms (A)
XX of importance in brewing which comprises treating a sample with at least
XX two primers (P1) that hybridize to a consensus region in the nucleic acid
XX of (A), at least part of the microbial nucleic acid is amplified, the
XX amplicon is treated with at least one probe (P2) that hybridizes
XX specifically with a sequence common to all (A) or specific for one or
XX more families, genera or species, and any formation of hybrids is
XX detected. The method is used to detect, identify and/or characterize
XX microorganisms in beer or brewing materials, particularly for detecting
XX contamination. The method may detect the entire range of contaminating

CC microbes, either as a general test for contamination or as a test
 CC specific for particular genera or (sub)species. It is quicker than known
 CC microbiological methods, and can detect several organisms in the same
 CC sample, including organisms not presently recognized as contaminants. The
 CC method provides an early indication of contamination and can be automated
 CC for high throughput analysis

XX Sequence 25 BP; 11 A; 2 C; 4 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 5; Length 25;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCTAG 25
 DB 1 TCGAGATATATGATATATCTAG 25

RESULT 2

AA61566 standard; DNA; 267 BP.

AC AAF61566;

DT 02-JUL-2001 (first entry)

DE Lactobacillus brevis 23S rRNA-spacer-5S rRNA DNA fragment.

KM 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

OS Lactobacillus brevis.

PN D819945964-A1.

PD 05-APR-2001.

PF 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PA (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K;

DR WPI; 2001-246136/26.

XX Detecting contaminating microorganisms in brewing, by nucleic acid
 PT amplification and hybridization, either non-specific or genus- or species
 PT -specific.

PS Claim 9(1); Page 9; 48pp; German.

XX This invention describes a novel method for detecting microorganisms (A)
 CC of importance in brewing which comprises treating a sample with at least
 CC two primers (P1) that hybridize to a consensus region in the nucleic acid
 CC of (A), at least part of the microbial nucleic acid is amplified, the
 CC amplicon is treated with at least one probe (P2) that hybridizes
 CC specifically with a sequence common to all (A) or specific for one or
 CC more families, genera or species, and any formation of hybrids is
 CC detected. The method is used to detect, identify and/or characterize
 CC microorganisms in beer or brewing materials, particularly for detecting
 CC contamination. The method may detect the entire range of contaminating
 CC microbes, either as a general test for contamination or as a test
 CC specific for particular genera or (sub)species. It is quicker than known
 CC microbiological methods, and can detect several organisms in the same
 CC sample, including organisms not presently recognized as contaminants. The
 CC method provides an early indication of contamination and can be automated
 CC for high throughput analysis

XX Sequence 267 BP; 77 A; 51 C; 77 G; 62 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 5; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGAGATATATGATATATCTAG 25
 DB 125 TCGAGATATATGATATATCTAG 149

RESULT 3

AB082965 standard; cDNA; 228 BP.

AC AB082965;

DT 15-JAN-2003 (first entry)

DE Human lung specific nucleic acid (LSNA) SEQ ID NO:108.

KM Human; lung specific gene; lung specific nucleic acid; LSG; LSNA; LSP;
 KM lung specific protein; lung cancer; vaccine; gene therapy; gene; ss.

OS Homo sapiens.

PN M020027236-A2.

PD 03-OCT-2002.

PF 26-OCT-2001; 2001.WO-US050385.

PR 26-OCT-2000; 2000US-0243459P.

PA (DIAD-) DIADEXUS INC.

PI Recipon H, Sun Y, Chen S, Liu C, Turner LR;

DR WPI; 2003-018930/01.

XX New isolated nucleic acid molecule, useful for treating lung cancer, and
 PT diagnosing or monitoring the presence of metastases of lung cancer in a
 PT patient.

PS Claim 1; Page 204; 262pp; English.

XX The present invention describes a isolated human lung specific nucleic
 CC acid (LSNA) molecule comprising a sequence selected from AB082858 to
 CC AB082993, which encodes a human lung specific protein (LSP) sequence
 CC selected from ABP54165 to ABP54272. Also described: (1) a method for
 CC determining the presence of a LSNA or LSP in a sample; (2) a vector
 CC comprising the novel nucleic acid; (3) a host cell comprising the vector
 CC of (2); (4) producing a polypeptide encoded by the novel nucleic acid;
 CC (5) an isolated polypeptide produced by the method of (4); (6) an
 CC antibody or its fragment that specifically binds to the polypeptide of
 CC (5); (7) diagnosing and monitoring the presence of metastases of lung
 CC cancer in a patient; (8) treating a patient with lung cancer; or (9) a
 CC vaccine comprising the polypeptide of (5) or the novel nucleic acid. The
 CC human LSNA and LSP sequences have cytosolic activity, and can be used in
 CC vaccines and gene therapy. The LSNA and LSP sequences can also be used
 CC for treating lung cancer, and diagnosing or monitoring the presence of
 CC metastases of lung cancer in a patient

XX Sequence 228 BP; 83 A; 29 C; 50 G; 66 T; 0 U; 0 Other;

Query Match 75.2%; Score 18.8; DB 8; Length 228;
 Best Local Similarity 90.9%; Pred. No. 6e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AGAATATGATATATCTAG 25
 DB 17 AGAATATGATATATCTAG 38

RESULT 4

AB082965/c standard; cDNA; 1324 BP.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 ; Search time 195.27 Seconds

(without alignments)
6054.391 Million cell updates/sec

Title: us-10-088-666-73

Perfect score: 25

Sequence: 1 tcgagatattgataatattctag 25

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	6	AX105801 Sequence
2	25	100.0	267	6	AX105729 Sequence
3	20.4	81.6	110000	2	AC141815 Apis mell
4	20.2	80.8	64701	2	AC079132 Homo sapi
5	20.2	80.8	178448	9	AC099805 Homo sapi
6	19.8	79.2	153768	2	AC113566 Canis fam
7	19.8	79.2	238427	2	AC113234 Canis fam
8	19.4	77.6	119113	5	EX005199 zebrafish
9	19.4	77.6	145011	2	CR548634 Danio rer
10	19.4	77.6	149163	2	AC073835 Homo sapi
11	19.4	77.6	154412	9	HS144P13 Human DNA
12	19.4	77.6	195130	9	AL355832 Human DNA
13	19.2	76.8	482	11	CR377191 Arabidops
14	19.2	76.8	1131	10	AF012235 Cryptomy
15	19.2	76.8	35291	5	EX649551 zebrafish
16	19.2	76.8	83537	5	EX004805 zebrafish
17	19.2	76.8	92624	8	AC003000 Arabidops
18	19.2	76.8	103029	9	AC133537 Homo sapi
19	19.2	76.8	123406	8	AC120508 Oryza sat

c	20	19.2	76.8	135129	2	AC148864 Canis fam
c	21	19.2	76.8	137923	5	EX640536 zebrafish
c	22	19.2	76.8	139257	2	AC150028 Canis fam
c	23	19.2	76.8	152841	2	AL929463 Danio rer
c	24	19.2	76.8	153438	2	AC117959 Rattus no
c	25	19.2	76.8	171061	2	EX950177 Danio rer
c	26	19.2	76.8	174846	2	AC141676 Apis mell
c	27	19.2	76.8	177722	2	EX928747 Danio rer
c	28	19.2	76.8	189219	9	AC092982 Homo sapi
c	29	19.2	76.8	189355	2	AC053539 Homo sapi
c	30	19.2	76.8	194717	2	AC113424 Homo sapi
c	31	19.2	76.8	202587	2	AC119862 Mus muscu
c	32	19.2	76.8	202922	2	EX901959 Danio rer
c	33	19.2	76.8	222615	5	AL929345 zebrafish
c	34	19.2	76.8	226614	2	EX957306 Danio rer
c	35	19.2	76.8	268324	5	AL954838 zebrafish
c	36	19.2	76.8	342116	2	AC130918 Rattus no
c	37	19.2	76.8	348465	2	AC131165 Rattus no
c	38	19	76.0	164921	2	AC150864 Bos tauru
c	39	19	76.0	226341	2	AC130020 Rattus no
c	40	18.8	75.2	228	6	AX566136 Sequence
c	41	18.8	75.2	787	3	AY573636 Drosophi
c	42	18.8	75.2	787	3	AY573639 Drosophi
c	43	18.8	75.2	788	3	AY573632 Drosophi
c	44	18.8	75.2	788	3	AY573633 Drosophi
c	45	18.8	75.2	788	3	AY573634 Drosophi

ALIGNMENTS

RESULT 1

LOCUS AX105801 25 bp DNA linear PAT 30-APR-2001

DEFINITION Sequence 73 from Patent WO0123605.

ACCESSION AX105801

VERSION AX105801.1 GI:13921814

KEYWORDS

SOURCE

ORGANISM

synthetic construct

synthetic construct

artificial sequences.

REFERENCE

AUTHORS Fandke, M., Gasch, A., and Berghof, K.

TITLE Method and nucleic acids for determining the presence of

JOURNAL micro-organisms specific to the brewing process

PATENT: WO 0123605-A 73 05-APR-2001;

Bioecon Diagnostics GmbH (DE)

FEATURES

source

1..25

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="spezifische Sequenz fur Lactobacillus brevis"

ORIGIN

Query Match 100.0% Score 25; DB 6; Length 25;

Best Local Similarity 100.0% Pred. No. 43;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 TCGAGATATTGAATATATCTAG 25

1 TCGAGATATTGAATATATCTAG 25

Db

RESULT 2

AX105729

LOCUS AX105729 267 bp DNA linear PAT 30-APR-2001

DEFINITION Sequence 1 from Patent WO0123605.

ACCESSION AX105729

VERSION AX105729.1 GI:13921742

KEYWORDS

SOURCE Lactobacillus brevis

ORGANISM Lactobacillus brevis

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source

Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 Lactobacillus.

1
 Fardke, M., Gasch, A. and Berghof, K.
 Method and nucleic acids for determining the presence of
 micro-organisms specific to the brewing process
 Patent: WO 0123605-A 1 05-Apr-2001;
 Biotecon Diagnostics GmbH (DE)
 Location/Qualifiers

1..267
 /organism="Lactobacillus brevis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1580"

ORIGIN

Query Match 100.0%; Score 25; DB 6; Length 267;
 Best local Similarity 100.0%; Pred. No. 25;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCGAGATATGATATATATCTAG 25
 125 TCGAGATATGATATATCTAG 149

Db

RESULT 3
 AC141815 0/c
 WCOMMENT
 Sequence split into 4 fragments LOCUS AC141815 Accession AC141815
 Fragment Name Begin End
 AC141815_0 1 110000
 AC141815_1 100001 210000
 AC141815_2 200001 310000
 AC141815_3 409619 496619
 LOCUS AC141815 496619 bp DNA linear HTG 19-MAR-2003
 DEFINITION Apis mellifera clone CH224-61C4, WORKING DRAFT SEQUENCE, 94
 AC141815
 unorderd pieces.
 AC141815
 AC141815.1 GI:29123999
 HTG; HTGS PHASE1; HTGS DRAFT.
 SOURCE
 ORGANISM
 Apis mellifera
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
 Apidae; Apis.
 1 (bases 1 to 409619)
 Muzny D.M., Adams C., Adio-Oduola B., Ali-osman, F.R., Allen, C.,
 Albrooks S.L., Amaralunga H.C., Are, J.R., Ayala, M., Banks, T.,
 Barbara, J., Benton, J., Blum, K., Blankenburg, K., Bonnin, D.,
 Boucek, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, U., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.R., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Hawlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,
 Homari, E., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korva, J., Kovar, C.,
 Kralovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, J., Louised, H.,
 Lozdo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Marindale, A., Martinez, E.,
 Massey, E., Mashiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Minter, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Nohbbat, K., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okunolu, G.,

ORAGUNE, N., OVIEDO, R., PACE, A., PAYTON, B., PEERY, J., PEREZ, L.,
 PETERS, L., PICKENS, R., PRIMUS, E., PU, L.L., QUILLES, M., REN, Y.,
 RIVES, M., ROJAS, A., ROJUBOKAN, I., RUIZ, M., RUIZ, S., SAVERY, G.,
 SCHERER, S., SCOTT, G., SHEN, H., SHOCHTARI, N., SISON, I.,
 SODERGREN, E., SONALKE, T., SPARKS, A., STANLEY, H., STONE, H.,
 SUTTON, A., SWALEK, A., TABOR, P., TAMERIS, A., TAMERIS, K., TANG, H.,
 TANEY, J., TAYLOR, C., TAYLOR, T., TELFORD, B., THOMAS, N., THOMAS, S.,
 UEMANI, K., VASQUEZ, L., VERA, V., VILLALON, D., VINSON, R., WANG, Q.,
 WANG, S., WARD-MOORE, S., WARREN, R., WASHINGTON, C., WATLINGTON, S.,
 WILLIAMS, G., WILLIAMSON, A., WIECZYK, R., WOODEN, S., WORLEY, K.,
 WU, C., WU, Y., WU, Y.F., ZHOU, J., ZORILLA, S., NELSON, D.,
 WEINSTOCK, G. and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 409619)
 Morley, K.C.
 Submitted (19-MAR-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.bgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: AMBL
 Center clone name: CH224-61C4
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 371769 bases at least Q40
 Consensus quality: 383951 bases at least Q30
 Consensus quality: 392893 bases at least Q20
 Estimated insert size: 375080; sum-of-coverage estimation
 Quality coverage: 3x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 94 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1189: contig of 1189 bp in length
 1190 1289: gap of unknown length
 1290 2420: contig of 1131 bp in length
 2421 2520: gap of unknown length
 2521 3713: contig of 1193 bp in length
 3714 3813: gap of unknown length
 3814 4963: contig of 1150 bp in length
 4964 5063: gap of unknown length
 5064 6387: contig of 134 bp in length
 6388 6487: gap of unknown length
 6488 7548: contig of 1061 bp in length
 7549 7648: gap of unknown length
 7649 9094: contig of 1446 bp in length
 9095 9194: gap of unknown length
 9195 10256: contig of 1062 bp in length
 10257 11389: gap of unknown length
 11390 11489: contig of 1033 bp in length
 11490 12816: gap of unknown length
 12817 12916: gap of 1327 bp in length
 12917 13951: contig of 1035 bp in length
 13952 14051: gap of unknown length
 14052 15422: contig of 1371 bp in length
 15423 15522: gap of unknown length

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 ; Search time 350.315 Seconds
(without alignments)
2184.416 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21

Sequence: 1 ccaagtcacacagctagttgt 21

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hrc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_est1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	85.7	779	8	BH542735
2	17.8	84.8	197	2	BF430206
3	17.8	84.8	204	2	BE477537
4	17.8	84.8	214	2	BE478105
5	17.8	84.8	218	2	BE486429
6	17.8	84.8	261	7	CR383046
7	17.8	84.8	290	7	CF614527
8	17.8	84.8	313	2	AW426583
9	17.8	84.8	337	5	BO640947
10	17.8	84.8	369	2	BE488687
11	17.8	84.8	387	2	BE475991
12	17.8	84.8	413	1	AV613476
13	17.8	84.8	418	7	CF362784
14	17.8	84.8	420	6	CH168898
15	17.8	84.8	421	2	BF230248
16	17.8	84.8	424	2	BE588467
17	17.8	84.8	427	1	AV618453
18	17.8	84.8	432	6	CH165949
19	17.8	84.8	442	2	BE476579
20	17.8	84.8	450	2	BE483083
21	17.8	84.8	451	4	BM432873
22	17.8	84.8	451	5	BP110601
23	17.8	84.8	454	6	CB221784
24	17.8	84.8	471	2	BE487090

25	17.8	84.8	472	2	BE481873	BE481873	167477	BA
26	17.8	84.8	477	1	AJ691187	AJ691187	170440	BA
27	17.8	84.8	477	2	BE483952	BE483952	170440	BA
28	17.8	84.8	479	4	BM431900	BM431900	168131	BA
29	17.8	84.8	493	2	BE483642	BE483642	169866	BA
30	17.8	84.8	494	4	BM432022	BM432022	169866	BA
31	17.8	84.8	495	5	BP107620	BP107620	169866	BA
32	17.8	84.8	498	2	BE481915	BE481915	169866	BA
33	17.8	84.8	502	2	BE485555	BE485555	172621	BA
34	17.8	84.8	509	2	BE482287	BE482287	168021	BA
35	17.8	84.8	512	2	BE481721	BE481721	167202	BA
36	17.8	84.8	512	2	BE487910	BE487910	177056	BA
37	17.8	84.8	513	2	BE476202	BE476202	158828	BA
38	17.8	84.8	513	2	BE480446	BE480446	165528	BA
39	17.8	84.8	515	1	AV667257	AV667257	827812	MA
40	17.8	84.8	515	7	CF361479	CF361479	827812	MA
41	17.8	84.8	519	4	BG680664	BG680664	329234	BA
42	17.8	84.8	522	2	BE481271	BE481271	166855	BA
43	17.8	84.8	524	2	BP110809	BP110809	166855	BA
44	17.8	84.8	525	4	BM431956	BM431956	166855	BA
45	17.8	84.8	534	1	AV606713	AV606713	166855	BA

ALIGNMENTS

RESULT 1
BH542735
LOCUS
DEFINITION
BOGXU95TR BOGX Brassica oleracea genomic clone BOGXU95, genomic survey sequence.
ACCESSION
BH542735
VERSION
BH542735.1 GI:17794516
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
1 (bases 1 to 779)
Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
JOURNAL
Other GSSs: BOGXU95TR
COMMENT
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location:Qualifiers
1..779
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BOGX"
/note="Vector: pGEM1, Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pGEM1 using BstXI linkers"

ORIGIN

Query Match 85.7%; Score 18; DB 8; Length 779;
Best local similarity 100.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGTCACACGAGTGT 19
DB 356 CCAAGTCACACGAGTGT 373

RESULT 2	197 bp	MRNA	linear	EST 29-NOV-2000
LOCUS	BF430206			
DEFINITION	2561771 MARC BSM Bos taurus cDNA 5', mRNA sequence.			
ACCESSION	BF430206			
VERSION	BF430206.1	GI:11442303		
KEYWORDS	EST.			
SOURCE	Bos taurus (cow)			
ORGANISM	Bos taurus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.			
AUTHORS	1 (bases 1 to 197)			
TITLE	Warren,W.C., Tao,N., Allison,T., Wagner,S., Mathalagan,N., Kata,S., Johnson,J., Smith,T.P.L. and Womack,J.			
JOURNAL	A survey of genes transcribed in bovine skeletal muscle			
COMMENT	Unpublished (2000)			
	Contact: Smith TPJ			
	USDA, ARS, US Meat Animal Research Center			
	PO Box 166, Clay Center, NE 68933-0166, USA			
	Tel: 402 762 4366			
	Fax: 402 762 4390			
	Email: smtthe@mail.marc.usda.gov			
	Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.			
	PCR Primers			
	FORWARD: GGAAACAGCTATGACCATG			
	BACKWARD: GTTTCCTCAGTCAGAC			
	Seq primer: AATTACCTCTCATTAAGGG.			
FEATURES	Location/Qualifiers			
SOURCE	1..197			
	/organism="Bos taurus"			
	/mol_type="mRNA"			
	/db_xref="taxon:9913"			
	/tissue_type="Skeletal muscle"			
	/lab_host="XLOLR"			
	/clone_lib="MARC BSM"			
	/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; Library obtained from stragene, catalog #93721. Library made from skeletal muscle of a two year old Holstein cow."			
ORIGIN				
Query Match	84.8%; Score 17.8; DB 2; Length 197;			
Best Local Similarity	90.5%; Pred. No. 3.2e-02;			
Matches	19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1 CCAAGTCAACACGATGTTGT 21			
	41 CCAATCTACACGATGTTGT 61			
Db				
RESULT 3				
LOCUS	BE477537	204 bp	mRNA	linear
DEFINITION	161590 BARC SBOV Bos taurus cDNA 5', mRNA sequence.			EST 27-MAR-2003
ACCESSION	BE477537			
VERSION	BE477537.1	GI:9597070		
KEYWORDS	EST.			
SOURCE	Bos taurus (cow)			
ORGANISM	Bos taurus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.			
	1 (bases 1 to 204)			
	Sontegard,T., Capucco,A.V., White,J., Van Tassel,C.P., Connor,E.B., Cho,J., Sultana,R., Shade,L., Wray,J.E., Wells,K.D. and Quackenbush,J.			
	Analysis of bovine mammary gland EST and functional annotation of the Bos taurus gene index			
	Mamm. Genome 13 (7), 373-379 (2002)			
	2215956			

COMMENT	12140684	Contact: Sonstegard TS
PUBMED	USDA, ARS, Beltsville Agricultural Research Center	
	Belg. 200 Km 2A, Beltsville, MD 20705, USA	
	Tel: 301 504 8416	
	Fax: 301 504 8414	
	Email: tads@psl.barc.usda.gov	
	Single pass sequencing. Bases called and alt. trimmed with phred	
	v0.980904.e. Vector identified by cross_match with the -minscore 18	
	and -mismatch 12 options.	
FEATURES	PCR Primers	
	FORWARD: AGGAACAGCTATGACCAT	
	BACKWARD: GTTTCACAGCAGC	
	Plate: 3 row: 7 column: 7	
	Seq primer: ATTTAGGTGACCTAATG.	
	Location/Qualifiers	
source	1..204	
	/organism="Bos taurus"	
	/mol_type="RNA"	
	/db_xref="taxon:9913"	
	/tissue_type="pooled"	
	/lab_host="DH10B"	
	/clone_idb="BARC SBOV"	
	/notes="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;	
	library made from pooled mRNA isolated from mammary	
	tissue at eight physiological, developmental, and disease	
	states."	
ORIGIN		
Query Match	84.8%; Score 17.8; DB 2; Length 204;	
Best Local Similarity	90.5%; Pred. No. 3.2e+02;	
Matches	19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
CY	1 CCAAGTCACACGCTAGTGT 21	
DB	138 CCAATCTACACGCTAGTGT 158	
RESULT 4		
LOCUS	BE478105	214 bp mRNA linear EST 27-MAR-2003
DEFINITION	162195 BARC SBOV Bos taurus cDNA 5', mRNA sequence.	
ACCESSION	BE478105	
VERSION	BE478105.1 GI:9597638	
KEYWORDS	EST.	
SOURCE	Bos taurus (cow)	
ORGANISM	Bos taurus	
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
	Bovinae; Bos.	
	1 (bases 1 to 214)	
REFERENCE	Sonstegard,T., Capuco,A.V., White,J., Van Tassel,C.P.,	
AUTHORS	Connor,E.B., Cho,J., Sultana,R., Shade,L., Wray,J.E., Wells,K.D.	
	and Quackenbush,J.	
	Analysis of bovine mammary gland EST and functional annotation of	
	the Bos taurus gene index	
	Mamm. Genome 13 (7), 373-379 (2002)	
TITLE	221335956	
JOURNAL	12140684	
MEDLINE	Contact: Sonstegard TS	
COMMENT	USDA, ARS, Beltsville Agricultural Research Center	
	Belg. 200 Km 2A, Beltsville, MD 20705, USA	
	Tel: 301 504 8416	
	Fax: 301 504 8414	
	Email: tads@psl.barc.usda.gov	
	Single pass sequencing. Bases called and alt. trimmed with phred	
	v0.980904.e. Vector identified by cross_match with the -minscore 18	
	and -mismatch 12 options.	
	PCR Primers	
	FORWARD: AGGAACAGCTATGACCAT	
	BACKWARD: GTTTCACAGCAGC	
	Plate: 6 row: A column: 2	
	Seq primer: ATTAGGTGACCTAATG.	

Tue Nov 16 16:09:54 2004

us-10-088-666-21.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:50:19 / Search time 36.2613 Seconds
(without alignments)
3127.876 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21

Sequence: 1 ccaagtcacacagtagtggt 21

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3625171 seqs, 2700493622 residues 7250342

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/FCI_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
19:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
21:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.8	84.8	207	US-09-960-352-4229	Sequence 4229, Ap
2	17.8	84.8	211	US-09-960-352-2061	Sequence 2061, Ap
3	17.8	84.8	229	US-09-960-352-4220	Sequence 4220, Ap
4	17.8	84.8	239	US-09-960-352-13451	Sequence 13451, A
5	17.8	84.8	242	US-09-960-352-2998	Sequence 2998, Ap
6	17.8	84.8	247	US-09-960-352-183	Sequence 183, App
7	17.8	84.8	292	US-09-960-352-7737	Sequence 7737, Ap
8	17.8	84.8	301	US-09-960-352-86	Sequence 86, App1
9	17.8	84.8	308	US-09-960-352-10392	Sequence 10392, A
10	17.8	84.8	340	US-09-960-352-2761	Sequence 2761, Ap
11	17.8	84.8	341	US-09-960-352-2191	Sequence 2191, Ap
12	17.8	84.8	342	US-09-960-352-11482	Sequence 11482, A

13	17.8	84.8	348	US-09-960-352-9403	Sequence 9403, Ap
14	17.8	84.8	358	US-09-960-352-8472	Sequence 8472, Ap
15	17.8	84.8	360	US-09-960-352-2079	Sequence 2079, Ap
16	17.8	84.8	360	US-09-960-352-12857	Sequence 12857, A
17	17.8	84.8	366	US-09-960-352-8659	Sequence 8659, Ap
18	17.8	84.8	366	US-09-960-352-14094	Sequence 14094, A
19	17.8	84.8	368	US-09-960-352-4223	Sequence 4223, Ap
20	17.8	84.8	378	US-09-960-352-2966	Sequence 2966, Ap
21	17.8	84.8	382	US-09-960-352-215	Sequence 215, App
22	17.8	84.8	385	US-09-960-352-10641	Sequence 10641, A
23	17.8	84.8	391	US-09-960-352-10595	Sequence 10595, A
24	17.8	84.8	395	US-09-960-352-2505	Sequence 2505, Ap
25	17.8	84.8	398	US-09-960-352-7013	Sequence 7013, Ap
26	17.8	84.8	399	US-09-960-352-9396	Sequence 9396, Ap
27	17.8	84.8	404	US-09-960-352-179	Sequence 179, App
28	17.8	84.8	404	US-09-960-352-7995	Sequence 2795, Ap
29	17.8	84.8	405	US-09-960-352-1246	Sequence 7246, Ap
30	17.8	84.8	406	US-09-960-352-12380	Sequence 12380, A
31	17.8	84.8	407	US-09-960-352-9846	Sequence 9846, Ap
32	17.8	84.8	416	US-09-960-352-11517	Sequence 11517, A
33	17.8	84.8	417	US-09-960-352-15053	Sequence 15053, A
34	17.8	84.8	419	US-09-960-352-7509	Sequence 7509, Ap
35	17.8	84.8	420	US-09-960-352-6117	Sequence 6117, Ap
36	17.8	84.8	427	US-09-960-352-7803	Sequence 7803, Ap
37	17.8	84.8	430	US-09-960-352-5911	Sequence 5911, Ap
38	17.8	84.8	458	US-09-960-352-9145	Sequence 9145, Ap
39	17.8	84.8	553	US-09-876-143-840	Sequence 840, App
40	17.8	84.8	586	US-09-876-143-140	Sequence 140, App
41	17.8	84.8	606	US-09-876-143-744	Sequence 744, App
42	17.8	84.8	641	US-09-876-143-197	Sequence 197, App
43	17.8	84.8	641	US-09-876-143-285	Sequence 285, App
44	17.8	84.8	690	US-09-876-143-791	Sequence 791, App
45	17.8	84.8	957	US-09-876-143-1436	Sequence 1436, Ap

ALIGNMENTS

RESULT 1
US-09-960-352-4229
Sequence 4229, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10296) C
CURRENT FILING DATE: 2001-09-24
CURRENT APPLICATION NUMBER: US/09/960,352
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 4229
LENGTH: 207
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 18-LIB34-062-Q1-E1-E5
US-09-960-352-4229

Query Match 84.8% Score 17.8; DB 9; Length 207;
Best Local Similarity 90.5%; Pred. No. 62;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAAGTCACACAGTAGTGT 21
Db 162 CCAATCTACACAGTAGTGT 182

RESULT 2
US-09-960-352-2061
Sequence 2061, Application US/09960352
Patent No. US20020137139A1

GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 2061
LENGTH: 211
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 09-LIB34-062-Q1-E1-C9
US-09-960-352-2061

Query Match 84.8%; Score 17.8; DB 9; Length 211;
Best Local Similarity 90.5%; Pred. No. 62;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 CCAAGTCAACAGCTAGTTGT 21
124 CCAATCTCAACGTAAGTTGT 144

RESULT 3
US-09-960-352-4220
Sequence 4220, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 4220
LENGTH: 229
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 18-LIB34-051-Q1-E1-B5
US-09-960-352-4220

Query Match 84.8%; Score 17.8; DB 9; Length 229;
Best Local Similarity 90.5%; Pred. No. 63;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 CCAAGTCAACAGCTAGTTGT 21
144 CCAATCTCAACGTAAGTTGT 164

RESULT 4
US-09-960-352-13451
Sequence 13451, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 13451
LENGTH: 239
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 57-LIB34-060-Q1-E1-G2
US-09-960-352-13451

Query Match 84.8%; Score 17.8; DB 9; Length 239;
Best Local Similarity 90.5%; Pred. No. 63;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 CCAAGTCAACAGCTAGTTGT 21
111 CCAATCTCAACGTAAGTTGT 131

RESULT 5
US-09-960-352-2998
Sequence 2998, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 2998
LENGTH: 242
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 13-LIB34-054-Q1-E1-D1
US-09-960-352-2998

Query Match 84.8%; Score 17.8; DB 9; Length 242;
Best Local Similarity 90.5%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 CCAAGTCAACAGCTAGTTGT 21
142 CCAATCTCAACGTAAGTTGT 162

RESULT 6
US-09-960-352-183
Sequence 183, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 183
LENGTH: 247
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 01-LIB34-051-Q1-E1-A9
US-09-960-352-183

Query Match 84.8%; Score 17.8; DB 9; Length 247;
Best Local Similarity 90.5%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:14 ; Search time 9.08108 Seconds
(without alignments)
1643.699 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21

Sequence: 1 ccaagtcacacagctagttgt 21

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*\n2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*\n3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*\n4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*\n5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*\n6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.2	77.1	525	US-09-248-796A-1527	Sequence 1527, Ap
2	16.2	77.1	963	US-09-328-352-3811	Sequence 3811, Ap
3	16.2	77.1	8878	US-08-759-444-2	Sequence 2, Appl
4	16.2	77.1	9880	US-08-680-897-1	Sequence 1, Appl
5	15.8	75.2	996	US-09-094-557-33	Sequence 33, Appl
6	15.8	75.2	1050	US-09-113-000C-2255	Sequence 2255, Ap
7	15.8	75.2	1053	US-09-071-035-409	Sequence 409, Appl
8	15.8	75.2	2666	US-09-524-101D-21	Sequence 21, Appl
9	15.4	73.3	780	US-09-248-796A-6461	Sequence 6461, Ap
10	15.4	73.3	1380	US-09-328-352-537	Sequence 537, Ap
11	15.2	72.4	735	US-09-248-796A-6839	Sequence 6839, Ap
12	15.2	72.4	1610	US-09-059-769-19	Sequence 19, Appl
13	15.2	72.4	2209	US-09-489-847-16	Sequence 16, Appl
14	15.2	72.4	3070	US-09-799-451-652	Sequence 652, Appl
15	15.2	72.4	13508	US-08-956-171E-120	Sequence 120, App
16	15.2	72.4	13508	US-08-956-171E-120	Sequence 120, App
17	14.8	70.5	36	US-08-612-895A-9	Sequence 9, Appl
18	14.8	70.5	36	US-09-093-293-9	Sequence 9, Appl
19	14.8	70.5	36	US-09-459-553-9	Sequence 9, Appl
20	14.8	70.5	36	US-09-882-246-9	Sequence 9, Appl
21	14.8	70.5	36	PCT-US94-10562A-9	Sequence 9, Appl
22	14.8	70.5	5749	US-09-603-757A-45	Sequence 45, Appl
23	14.6	69.5	105	US-09-206-942-9	Sequence 9, Appl
24	14.6	69.5	127	US-08-956-171E-4780	Sequence 4780, Ap
25	14.6	69.5	127	US-08-781-986A-4780	Sequence 4780, Ap
26	14.6	69.5	267	US-09-253-991A-6214	Sequence 6214, Ap
27	14.6	69.5	300	US-08-956-171E-4419	Sequence 4419, Ap

28	14.6	69.5	300	US-08-781-986A-4419	Sequence 4419, Ap
29	14.6	69.5	302	US-08-956-171E-4448	Sequence 4448, Ap
30	14.6	69.5	302	US-08-781-986A-4448	Sequence 4448, Ap
31	14.6	69.5	355	US-08-956-171E-4233	Sequence 4233, Ap
32	14.6	69.5	355	US-08-781-986A-4233	Sequence 4233, Ap
33	14.6	69.5	359	US-08-956-171E-4081	Sequence 4081, Ap
34	14.6	69.5	359	US-08-781-986A-4081	Sequence 4081, Ap
35	14.6	69.5	361	US-08-956-171E-4071	Sequence 4071, Ap
36	14.6	69.5	361	US-08-781-986A-4071	Sequence 4071, Ap
37	14.6	69.5	400	US-08-956-171E-3652	Sequence 3652, Ap
38	14.6	69.5	400	US-08-781-986A-3652	Sequence 3652, Ap
39	14.6	69.5	400	US-08-956-171E-3756	Sequence 3756, Ap
40	14.6	69.5	400	US-08-956-171E-3843	Sequence 3843, Ap
41	14.6	69.5	400	US-08-956-171E-4470	Sequence 4470, Ap
42	14.6	69.5	400	US-08-781-986A-3647	Sequence 3647, Ap
43	14.6	69.5	400	US-08-781-986A-3652	Sequence 3652, Ap
44	14.6	69.5	400	US-08-781-986A-3756	Sequence 3756, Ap
45	14.6	69.5	400	US-08-781-986A-3843	Sequence 3843, Ap

ALIGNMENTS

RESULT 1
US-09-248-796A-1527/c
Sequence 1527, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248, 796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074, 725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096, 409
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 1527
LENGTH: 525
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-1527

Query Match 77.1%, Score 16.2, DB 4, Length 525;
Best Local Similarity 85.7%, Pred. No. 62;
Matches 18, Conservative 0, Mismatches 3, Indels 0, Gaps 0;

QY 1 CCAAGTCACACAGCTAGTTGT 21
Db 279 CCAAGTCACACAGCTAGTTT 259

RESULT 2

US-09-328-352-3811
Sequence 3811, Application US/09328352
Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

Query Match 77.1%, Score 16.2, DB 4, Length 963;

Best Local Similarity 85.7%; Pred. No. 69;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCAAGTCACACGAGTGTGT 21
DB 382 CAAAGACACACGAGTGTAT 402

RESULT 3

US-08-759-444-2/c
Sequence 2, Application US/08759444
Patent No. 5824309
GENERAL INFORMATION:
APPLICANT: Dassarma, Shiladitya
APPLICANT: Moished, Fazeela
APPLICANT: Stuart, Elizabeth
APPLICANT: Black, Samuel
TITLE OF INVENTION: RECOMBINANT GAS VESICLES AND USES THEREOF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,444
FILING DATE: 05-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/008,200
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07880/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 8878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-759-444-2

Query Match 77.1%; Score 16.2; DB 1; Length 8878;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCAAGTCACACGAGTGTGT 21
DB 6765 CCAAGTCACACGAGTGTGT 6745

RESULT 4
US-08-680-897-1/c
Sequence 1, Application US/08680897
Patent No. 6008051

GENERAL INFORMATION:
APPLICANT: Dassarma, Shiladitya
APPLICANT: Halladay, John
APPLICANT: Ng, Wai-lap
TITLE OF INVENTION: RECOMBINANT VECTOR AND PROCESS FOR CELL
TITLE OF INVENTION: FLUTATION
NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,897
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/271,270
FILING DATE:
APPLICATION NUMBER: 944,581
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiclic, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8680
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9880 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-680-897-1

Query Match 77.1%; Score 16.2; DB 3; Length 9880;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCAAGTCACACGAGTGTGT 21
DB 7765 CCAAGTCACACGAGTGTGT 7745

RESULT 5

US-09-094-557-33
Sequence 33, Application US/09094557
Patent No. 6225531
GENERAL INFORMATION:
APPLICANT: Kakitani, Makoto
APPLICANT: Umemoto, Naoyuki
APPLICANT: Ishida, Isao
APPLICANT: Imamatsu, Akihiro
APPLICANT: Yoshikawa, Masaki
APPLICANT: Yamaoka, Naoto
TITLE OF INVENTION: GLUCAN ELICITOR RECEPTOR, DNA MOLECULE
TITLE OF INVENTION: CODING THEREFOR, FUNGUS-RESISTANT PLANTS TRANSFORMED WITH
TITLE OF INVENTION: THE DNA MOLECULE AND METHOD FOR CREATING THE PLANTS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 ; Search time 40.7387 Seconds
(without alignments)
2705.971 Million cell updates/sec

Title: US-10-088-666-21

Perfect score: 21

Sequence: 1 ccaagtcacacagtagtgc 21

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003s:*
7: Geneseq2004s:*
8: Geneseq2005s:*
9: Geneseq2006s:*
10: Geneseq2007s:*
11: Geneseq2008s:*
12: Geneseq2009s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	5	AAF61586 Lactobac
2	21	100.0	267	5	AAF61566 Lactobac
3	18	85.7	2215	3	AAZ43884 C. Victoria
4	17.8	84.8	207	8	ABX39064 Bovine ES
5	17.8	84.8	211	8	ABX36896 Bovine ES
6	17.8	84.8	229	8	ABX39055 Bovine ES
7	17.8	84.8	239	8	ABX48286 Bovine ES
8	17.8	84.8	242	8	ABX37833 Bovine ES
9	17.8	84.8	247	8	ABX35018 Bovine ES
10	17.8	84.8	292	8	ABX42572 Bovine ES
11	17.8	84.8	301	8	ABX34921 Bovine ES
12	17.8	84.8	308	8	ABX45227 Bovine ES
13	17.8	84.8	340	8	ABX37596 Bovine ES
14	17.8	84.8	341	8	ABX37026 Bovine ES
15	17.8	84.8	342	8	ABX46317 Bovine ES
16	17.8	84.8	348	8	ABX44238 Bovine ES
17	17.8	84.8	358	8	ABX43307 Bovine ES
18	17.8	84.8	360	8	ABX35914 Bovine ES
19	17.8	84.8	360	8	ABX47692 Bovine ES
20	17.8	84.8	366	8	ABX43494 Bovine ES
21	17.8	84.8	366	8	ABX48929 Bovine ES

22	17.8	84.8	368	8	ABX39058	Abx39058 Bovine ES
23	17.8	84.8	378	8	ABX37801	Abx37801 Bovine ES
24	17.8	84.8	382	8	ABX35050	Abx35050 Bovine ES
25	17.8	84.8	385	8	ABX45476	Abx45476 Bovine ES
26	17.8	84.8	391	8	ABX45430	Abx45430 Bovine ES
27	17.8	84.8	395	8	ABX37340	Abx37340 Bovine ES
28	17.8	84.8	398	8	ABX41848	Abx41848 Bovine ES
29	17.8	84.8	399	8	ABX44231	Abx44231 Bovine ES
30	17.8	84.8	404	8	ABX35014	Abx35014 Bovine ES
31	17.8	84.8	404	8	ABX37630	Abx37630 Bovine ES
32	17.8	84.8	405	8	ABX42081	Abx42081 Bovine ES
33	17.8	84.8	405	8	ABX47215	Abx47215 Bovine ES
34	17.8	84.8	407	8	ABX44681	Abx44681 Bovine ES
35	17.8	84.8	416	8	ABX46352	Abx46352 Bovine ES
36	17.8	84.8	417	8	ABX44888	Abx44888 Bovine ES
37	17.8	84.8	419	8	ABX42344	Abx42344 Bovine ES
38	17.8	84.8	420	8	ABX40952	Abx40952 Bovine ES
39	17.8	84.8	427	8	ABX42638	Abx42638 Bovine ES
40	17.8	84.8	430	8	ABX40746	Abx40746 Bovine ES
41	17.8	84.8	458	8	ABX39980	Abx39980 Bovine ES
42	17.8	84.8	553	6	ABN73905	Abn73905 Bovine em
43	17.8	84.8	586	6	ABN73190	Abn73190 Bovine em
44	17.8	84.8	606	6	ABN73809	Abn73809 Bovine em
45	17.8	84.8	641	6	ABN73340	Abn73340 Bovine em

ALIGNMENTS

RESULT 1
AAF61586
ID AAF61586 standard; DNA; 21 BP.
XX
XX
AC AAF61586;
XX
DT 02-JUL-2001 (first entry)
XX
DE Lactobacillus brevis 23S rRNA/5S rRNA specific DNA probe.
XX
XX 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.
XX
OS Lactobacillus brevis.
XX
PN DE19945964-A1.
PD 05-APR-2001.
XX
PF 24-SEP-1999; 99DE-01045964.
XX
PR 24-SEP-1999; 99DE-01045964.
XX
PA (BIOT-) BIOTRECON DIAGNOSTICS GMBH.
PI Fandke M, Gasch A, Berghof K;
XX WPI, 2001-246136/26.
XX
PT Detecting contaminating microorganisms in brewing, by nucleic acid
PT amplification and hybridization, either non-specific or genus- or species
PT -specific.
XX
XX Claim 9(i); Page 15; 48pp; German.
XX
XX This invention describes a novel method for detecting microorganisms (A)
XX of importance in brewing which comprises treating a sample with at least
XX two primers (P1) that hybridize to a consensus region in the nucleic acid
XX of (A), at least part of the microbial nucleic acid is amplified, the
XX amplicon is treated with at least one probe (P2) that hybridizes
XX specifically with a sequence common to all (A) or specific for one or
XX more families, genera or species, and any formation of hybrids is
XX detected. The method is used to detect, identify and/or characterize
XX microorganisms in beer or brewing materials, particularly for detecting
XX contamination. The method may detect the entire range of contaminating

CC microbes, either as a general test for contamination or as a test
 CC specific for particular genera or (sub)species. It is quicker than known
 CC microbiological methods, and can detect several organisms in the same
 CC sample, including organisms not presently recognized as contaminants. The
 CC method provides an early indication of contamination and can be automated
 CC for high throughput analysis

XX Sequence 21 BP; 7 A; 5 C; 4 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 5; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAAGTCACACGCTAGTGT 21
 1 CCAAGTCACACGCTAGTGT 21

Db 1 CCAAGTCACACGCTAGTGT 21

RESULT 2

AA61566

AA61566 standard; DNA; 267 BP.

AC AA61566;

DT 02-JUL-2001 (first entry)

DE Lactobacillus brevis 23S rRNA-spacer-5S rRNA DNA fragment.

KM 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; ss.

OS Lactobacillus brevis.

PN DE19945964-A1.

PD 05-APR-2001.

PF 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PA (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K;

DR WPI; 2001-246136/26.

PT Detecting contaminating microorganisms in brewing, by nucleic acid
 PT amplification and hybridization, either non-specific or genus- or species
 PT -specific.

PS Claim 9(1); Page 9; 48pp; German.

XX This invention describes a novel method for detecting microorganisms (A)
 CC of importance in brewing which comprises treating a sample with at least
 CC two primers (P1) that hybridize to a consensus region in the nucleic acid
 CC of (A), at least part of the microbial nucleic acid is amplified, the
 CC amplicon is treated with at least one probe (P2) that hybridizes
 CC specifically with a sequence common to all (A) or specific for one or
 CC more families, genera or species, and any formation of hybrids is
 CC detected. The method is used to detect, identify and/or characterize
 CC microorganisms in beer or brewing materials, particularly for detecting
 CC contamination. The method may detect the entire range of contaminating
 CC microbes, either as a general test for contamination or as a test
 CC specific for particular genera or (sub)species. It is quicker than known
 CC microbiological methods, and can detect several organisms in the same
 CC sample, including organisms not presently recognized as contaminants. The
 CC method provides an early indication of contamination and can be automated
 CC for high throughput analysis

XX Sequence 267 BP; 77 A; 51 C; 77 G; 62 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 5; Length 267;
 Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAAGTCACACGCTAGTGT 21

Db 103 CCAAGTCACACGCTAGTGT 123

RESULT 3

AAZ43864/C

AAZ43864 standard; cDNA; 2215 BP.

AC AAZ43864;

DT 10-MAR-2000 (first entry)

DE C. vicina LSP-2 cDNA.

KM Arylphorin; calliphorin; ABP; receptor-binding domain; plant protection;

KM hygiene; veterinary medicine; LSP-2; ds.

OS Calliphora vicina.

PN Key Location/Qualifiers

FT CDS 16..2121

FT /tag= a

FT /product= "LSP-2"

PN DE19824492-A1.

PD 09-DEC-1999.

PF 02-JUN-1998; 98DE-01024492.

PR 02-JUN-1998; 98DE-01024492.

PA (FARB) BAYER AG.

PI Scheller K, Hansen I, Gutschmann V, Turberg A;

DR WPI; 2000-054341/05.

DR P-PSDB; AAY51007.

PT Receptor-binding domain useful for identifying new substances for plant

PT protection, hygiene or veterinary medicine.

PS Example A; Page 40-43; 50pp; German.

XX This invention describes novel receptor-binding domain of the Calliphora
 CC vicina arylphorin protein (also known as calliphorin). The receptor-
 CC binding domain, nucleic acid constructs, host cells and antibodies
 CC described in the invention are useful for discovering new active
 CC substances for plant protection, hygiene or veterinary medicine, in
 CC particular for combination, which influences the interaction between
 CC arylphorin and the arylphorin receptor. This sequence encodes the C.
 CC vicina LSP-2 protein

XX Sequence 2215 BP; 634 A; 524 C; 433 G; 624 T; 0 U; 0 Other;

Query Match 85.7%; Score 18; DB 3; Length 2215;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAAGTCACACGCTAGT 18
 1687 CCAAGTCACACGCTAGT 1670

Db

RESULT 4

ABX39064

ABX39064 standard; cDNA; 207 BP.

AC ABX39064;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 ; Search time 4454.01 Seconds
(without alignments)
2184.416 Million cell updates/sec

Title: US-10-088-666-1

Perfect score: 267

Sequence: 1 tatatggaagcraagaccctc.....agtagtcggggatcgcccc 267

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 1821985908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est1:*
9: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163.4	61.2	793	B2369094	Cot100.1.1
2	105.4	39.5	3268	BH770998	LMGtag72
3	105.4	39.5	6499	BH771024	LMGtag74
4	58.2	21.8	342	CR469421	CR469421
5	58.2	21.8	690	BH687637	BH687637
6	51.6	19.3	1258	CR477397	CR477397
7	50.4	18.9	391	BH770957	LMGtag68
8	50.2	18.8	607	AO990586	RfC0182
9	50.2	18.8	393	AO989492	RfC00040
10	49.2	18.4	453	CC144308	B3 Neisse
11	48.6	18.2	833	BU003208	CG34404
12	48.6	18.2	833	CG659016	PR10132
13	48.6	18.2	844	CG664279	PR10146C
14	48.6	18.2	872	CG662374	PR10141b
15	48.6	18.2	879	CG667131	PR10154b
16	47.2	17.7	255	BH770635	LMGtag39
17	47.2	17.7	333	CR470260	CR470260
18	47.2	17.7	698	AG613316	Beschertich
19	46	17.2	534	CG681720	PR10131d
20	46	17.2	664	AG613279	Beschertich
21	46	17.2	735	CG667389	PR10155a
22	45.8	17.2	544	AO989631	RfC00197
23	45.8	17.2	655	AO990982	RfC01832
24	44.8	16.8	1073	AF114216	AF114216

c 25	44.4	16.6	646	9	CL660916	PR10138b
c 26	44.4	16.6	685	9	CL656043	PR10125c
c 27	44.4	16.6	786	9	CL654528	PR10120d
c 28	44.4	16.6	811	9	CL673498	PR1019c
c 29	44.4	16.6	820	9	CL668272	PR10157b
c 30	44.2	16.6	643	2	AM948147	RCO-MT001
c 31	44	16.5	381	1	AU180248	AU180248
c 32	44	16.5	814	7	CNS86292	USDA-PP 1
c 33	43.8	16.4	1101	8	AF029514	AF029514
c 34	43.6	16.3	351	1	AI903023	AI903023
c 35	43.2	16.2	1197	4	B2579062	mb2 6160
c 36	43.2	16.1	343	4	B1895648	ESTED01
c 37	42.6	16.0	330	2	BE092318	IL2-BT073
c 38	42.6	16.0	694	2	AM948136	RCO-MT001
c 39	42.6	16.0	730	2	AM948130	RCO-MT001
c 40	42.6	16.0	738	2	AM948138	RCO-MT001
c 41	42.4	15.9	158	2	AM858409	CMO-CT034
c 42	42.4	15.9	282	7	CN025635	UNC-P4C14
c 43	42.4	15.9	518	5	BX548956	BX548956
c 44	42.4	15.9	548	5	BX548957	BX548957
c 45	42.4	15.9	791	9	CL670555	PR10162b

ALIGNMENTS

RESULT 1
B2369094
LOCUS
DEFINITION
Cot100.1.2.D04 Maize Cot100 library Zea mays genomic, genomic
survey sequence.

ACCESSION
B2369094
VERSION
B2369094.1
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
Yuan, Y., Sammlinger, P. and Bennetzen, J. L.
TITLE
High Cot sequence analysis of the maize genome
JOURNAL
Unpublished (2003)
COMMENT
Contact: Bennetzen JL
Department of Biological Sciences
Purdue University
Hansen 339f, Purdue University, West Lafayette, IN 47907, USA
Tel: 765 494 4919
Fax: 765 496 1496
Email: maize@bilbo.bio.purdue.edu

Forward and reverse reads were assembled when significant overlap was detected.
Seq primer: T7 and T3
Class: shotgun.
Location/Qualifiers
1. 793
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/dev_stage="10 days seedling"
/lab_host="DH10B"
/clone_lib="Maize Cot100 library"
/note="Organ: young leaves; Vector: pcr4TOPO; Maize genomic DNA was sheared to fragments averaging about 1.8 kb, was denatured and then reassociated in 1 x SSC at 65°C. After a given Cot value was reached, aliquots were run over a hydroxyapatite (HAP) column in order to separate single stranded DNA from double stranded DNA. The single stranded DNA was then converted to a double stranded form with one round of Klenow DNA polymerase treatment with random 9-mer primers. The double-stranded fragments were then further size-selected over an agarose

FEATURES
source

gel and cloned into the PCR4 TOPO vector."

ORIGIN

Query Match 51.2%; Score 163.4; DB 8; Length 793;
Best Local Similarity 81.1%; Pred. No. 1.3e-40;
Matches 215; Conservative 0; Mismatches 46; Indels 4; Gaps 2;

3 TATGAAGTAAGACCCCTGAGAGATGATCAGTAGATAGGCTGGAAGTAGACAGCCCGTG 62
Db TTTGAGAGTAAGACCCCAAGAGATATGTGTAGATAGAGATGGAAGTAGAGTGGGTG 191
Cy AGGGGTGAGACCGACCACTATATGCTGTGAGAGACTTAAACCAAGTCAACAGTAGTTG 122
Db AGGATGAGACCGACCACTATATGCTGTGAGAGACTTAAACCAAGTCAAGTGTGAGAGTT 251
Cy TTTGAGATATATGATATATCTATGTTTGGAGAGAGAGATCTCTATAGTGTG 182
Db CTTCATG--TGAAAGATATGTTTATGTTTGAACACAG--CTTCAAGGTGTG 307
Cy 183 GCGATGACCTGAGAGATACCTGTTCCATGCGGACACAGAGTTAGCTTCGACG 242
Db 308 ATGATGACCTGAGAGATACCTGTTCCATGCGGACACAGAGTTAGCTTCGACG 367
Cy 243 CCGATAGTAGTTGGGGATGCCCC 267
Db 368 CCAAGTAGTTGGGGATGCCCC 392

RESULT 2
BH770998 3268 bp DNA linear GSS 01-MAY-2002
LOCUS LIMGtag721 MG1363 Random Sequence Tag Library Lactococcus lactis
DEFINITION subsp. cremoris genomic, genomic survey sequence.
ACCESSION BH770998
VERSION BH770998.1 GI:20373955
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
AUTHORS 1. (bases 1 to 3268)
TITLE Bolotin, A., Ehrlich, S.D. and Sorokin, A.
JOURNAL Studies of genomes of dairy bacteria Lactococcus lactis
COMMENT Sci. Aliments (2002) In press
Contact: Sorokin A
Genetique Microbiome

FEATURES
source
1. 3268
Location/Qualifiers
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MG1363"
/sub_species="cremoris"
/db_xref="taxon:1359"
/clone_11b="MG1363 Random Sequence Tag Library"
/note="Vector: pSGM2; Site 1: SmaI; Library of
chromosomal fragments of L. lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 39.5%; Score 105.4; DB 8; Length 3268;
Best Local Similarity 67.3%; Pred. No. 4.6e-22;
Matches 185; Conservative 0; Mismatches 76; Indels 14; Gaps 2;

7 GAAGTAAGACCCCTGAGAGATGATCAGTAGATAGGCTGGAAGTAGACAGCCCGTAGGC 66

Db 361 GAATTAAGACCCCAAGAGATATCTGTAGATAGGCTGGAAGTAGAGTTCGAGAC 420
Cy 67 GTGAGACGACCACTATATGCTGTGAGAGACTTAACCA-----GTCAACAC 115
Db 421 TTGAGACGACCACTATATGCTGTGAGAGACTTAACCAAGTCAATATAATATGC 480
Cy 116 GTAGTTGTTGCAATATATGATATATCTATGTTTGGAGAGAGAGTCTCTATAG 175
Db 481 TTATGTTTGAAGTAGAGATATGTTTATGTTTGAATGTTCAAGTACATTTTA 540
Cy 176 TG---TGTTGGCATAGCTGAGATACCTGTTCCATGCGGACACAGAGTTAAG 232
Db 541 TGATTTGCATCATTCATGATGATGATACCTGTTCCATGCGGACACAGAGTTAAG 600
Cy 233 CTTCAGACGCGGATAGTAGTTGGGGATGCCCC 267
Db 601 TCATCTACGCGGAGATCTTGGGGGTTGCCCC 635

RESULT 3
BH771024 6499 bp DNA linear GSS 01-MAY-2002
LOCUS LIMGtag746 MG1363 Random Sequence Tag Library Lactococcus lactis
DEFINITION subsp. cremoris genomic, genomic survey sequence.
ACCESSION BH771024
VERSION BH771024.1 GI:20373981
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
AUTHORS 1. (bases 1 to 6499)
TITLE Bolotin, A., Ehrlich, S.D. and Sorokin, A.
JOURNAL Studies of genomes of dairy bacteria Lactococcus lactis
COMMENT Sci. Aliments (2002) In press
Contact: Sorokin A
Genetique Microbiome

FEATURES
source
1. 6499
Location/Qualifiers
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MG1363"
/sub_species="cremoris"
/db_xref="taxon:1359"
/clone_11b="MG1363 Random Sequence Tag Library"
/note="Vector: pSGM2; Site 1: SmaI; Library of
chromosomal fragments of L. lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 39.5%; Score 105.4; DB 8; Length 6499;
Best Local Similarity 67.3%; Pred. No. 5.3e-22;
Matches 185; Conservative 0; Mismatches 76; Indels 14; Gaps 2;

7 GAAGTAAGACCCCTGAGAGATGATCAGTAGATAGGCTGGAAGTAGACAGCCCGTAGGC 66
Db 5177 GAATTAAGACCCCAAGAGATATCTGTAGATAGGCTGGAAGTAGAGTTCGAGAC 5236
Cy 67 GTGAGACGACCACTATATGCTGTGAGAGACTTAACCA-----GTCAACAC 115
Db 5237 TTGAGACGACCACTATATGCTGTGAGAGACTTAACCAAGTCAATATAATATGC 5296
Cy 116 GTAGTTGTTGCAATATATGATATATCTATGTTTGGAGAGAGAGTCTCTATAG 175

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:50:19 / Search time 461.036 Seconds

(without alignments)
3127.876 Million cell updates/sec

Title: US-10-088-666-1

Perfect score: 267

Sequence: 1 tatatggaagaaagaccctc.....agtatgtggggatcgcccc 267

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3625171 seqs, 2700493622 residues

Total number of hits satisfying chosen parameters: 7250342

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications NA: *
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq: *
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq: *
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115.8	43.4	31702	US-10-398-221-3	Sequence 3, Appl1
2	115.6	43.3	3510	US-08-961-527-53	Sequence 53, Appl1
3	115.6	43.3	3510	US-10-158-844-53	Sequence 359, Appl1
4	115.6	43.3	3989	US-09-070-927A-359	Sequence 359, Appl1
5	115.6	43.3	5048	US-09-884-465A-2	Sequence 2, Appl1
6	115.6	43.3	8411	US-08-961-527-16	Sequence 16, Appl1
7	115.6	43.3	8411	US-10-158-844-16	Sequence 16, Appl1
8	115.6	43.3	11915	US-08-961-527-96	Sequence 96, Appl1
9	115.6	43.3	11915	US-10-158-844-96	Sequence 96, Appl1
10	114.2	42.8	1389	US-10-398-221-318	Sequence 318, Ap
11	114.2	42.8	2631	US-10-398-221-21	Sequence 2, Appl1
12	114.2	42.8	5349	US-10-398-221-3795	Sequence 3795, Ap

13	114.2	42.8	319630	US-10-398-221-7	Sequence 7, Appl1
14	114.2	42.8	684707	US-10-398-221-9	Sequence 9, Appl1
15	114.2	42.8	684707	US-10-398-221-9	Sequence 9, Appl1
16	114.2	42.8	1163020	US-10-398-221-10	Sequence 10, Appl1
17	114.2	42.8	3011208	US-10-398-221-2058	Sequence 2058, Ap
18	114.2	42.8	3011208	US-10-398-221-2058	Sequence 2058, Ap
19	114	42.7	9797	US-09-070-927A-550	Sequence 550, App
20	114	42.7	22960	US-09-070-927A-345	Sequence 345, App
21	113.2	42.4	882	US-10-398-221-3395	Sequence 3395, App
22	113.2	42.4	1907	US-10-398-221-1847	Sequence 1847, App
23	113.2	42.4	4199	US-10-398-221-3890	Sequence 3890, App
24	113.2	42.4	6525	US-10-398-221-2041	Sequence 2041, App
25	111.6	41.8	840	US-08-781-986A-508	Sequence 508, App
26	111.6	41.8	840	US-10-329-624-508	Sequence 508, App
27	109.6	41.0	400	US-08-781-986A-3738	Sequence 3738, App
28	109.6	41.0	400	US-10-329-624-3738	Sequence 3738, App
29	106	39.7	400	US-08-781-986A-3624	Sequence 3624, App
30	106	39.7	400	US-10-329-624-3624	Sequence 3624, App
31	104.4	39.1	386	US-08-781-986A-4064	Sequence 4064, App
32	104.4	39.1	386	US-10-329-624-4064	Sequence 4064, App
33	101.2	37.9	6591	US-08-781-986A-3114	Sequence 3114, App
34	101.2	37.9	6591	US-10-329-624-3114	Sequence 3114, App
35	100.8	37.8	400	US-08-781-986A-3768	Sequence 3768, App
36	100.8	37.8	400	US-10-329-624-3768	Sequence 3768, App
37	100.6	37.7	400	US-08-781-986A-3638	Sequence 3638, App
38	100.6	37.7	400	US-08-781-986A-3748	Sequence 3748, App
39	100.6	37.7	400	US-10-329-624-3748	Sequence 3748, App
40	100.6	37.7	400	US-10-329-624-3748	Sequence 3748, App
41	100.6	37.7	458	US-08-781-986A-3757	Sequence 3757, App
42	100.6	37.7	458	US-10-329-624-3757	Sequence 3757, App
43	100.6	37.7	15249	US-08-781-986A-102	Sequence 102, App
44	100.6	37.7	15249	US-10-329-624-102	Sequence 102, App
45	100.6	37.7	30246	US-08-781-986A-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1
US-10-398-221-3/c
Sequence 3, Application US/10398221
Publication No. US20040018514A1
GENERAL INFORMATION:
APPLICANT: KUNST, Frederick
TITLE OF INVENTION: Listeria innocua, genome and applications
FILE REFERENCE: 344 702 - US
CURRENT APPLICATION NUMBER: US/10/398, 221
CURRENT FILING DATE: 2003-03-27
PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
PRIOR FILING DATE: 2001-10-04
PRIOR FILING DATE: 2000-10-04
NUMBER OF SEQ ID NOS: 4025
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 31702
TYPE: DNA
ORGANISM: Listeria innocua
US-10-398-221-3

Query Match 43.4%; Score 115.8; DB 16; Length 31702;
Best Local Similarity 69.7%; Pred. No. 2.3e-26;
Matches 189; Conservative 0; Mismatches 72; Indels 10; Gaps 2;

QY 6 GGAAGTAAGACCCCTGAGAGATGATCAGGTAGTACGCTGGAAGTACGAGCCGTCGAG 65
DB 30606 GAAAGTAAGATCCCTGAAAGATATATAGGTAGTATGAGTGGCAGGTGAGGATA 30547
QY 66 CGTGGAGCGGACCAAGTACTGATGCTGCTGAGAGACTTAACCAAGTCAACAGTATTGTTT 125
DB 30546 CATGGAGCGGACCAAGTACTGATGCTGAGAGACTTAACCAAGTCAACAGTATTGTTT 30488

QY 126 CGAGATATTAATTAATATCTAGTTTGAAGGAGAAGTTCT-----CTTATAGT 176
DB 30487 CTAACGTACACCTTCTTCTCTAGTTTGAAGAGCATCTTTCAACACTGCATATGT 30428
QY 177 GTGTGGCGATAGCCTGAAGATACACTGTTCCCAAGCCGAACAGAAATTAACTTC 236
DB 30427 CTGTAGTATAGGCGAAGAGGTACACCCGTTCCATCCCGAACAGCGTTAAAGCTTC 30368
QY 237 AGCAGCGCGATAGTATGGGGGATCGCC 267
DB 30367 TCTGCGCCATGTAGTATGGGGGCTTCC 30337

RESULT 2

US-08-961-527-53/c
; Sequence 53, Application US/08961527
; Publication No. US20020032323A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3510 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-53

Query Match 43.3%; Score 115.6; DB 8; Length 3510;
Best Local Similarity 66.8%; Pred. No. 1.1e-26;
Matches 183; Conservative 0; Mismatches 84; Indels 7; Gaps 1;

QY 1 TATATGGAAGTAAAGCCCTTGAGAGATGATCAGGTAGCTAGGAGTGAAGTGAAGCCCG 60
DB 3288 TATATATCAGTAGAGCCCTTGAGAGATGATCAGGTAGCTAGGAGTGAAGTGAAGTGG 3229
QY 61 TGAGCGGTGAGAGCGGACGATCTACTATCGTGAAGACTTAACCAAGTCAACAGTAGT 120
DB 3228 CGACACATGTAGCGGACGATCTACTATCGTGAAGACTTAACCAAGTCAACAGTAGT 3169
QY 121 TGTTCGAGATATTAATTAATATCTAGTTTGAAGGAGAG-----TTCCTTAT 173
DB 3168 TGAAGCGGACGCTTTCTTAATTAATGAATAGATTAATTTTGAAGTATTAAGTACG 3109
QY 174 AGTGTGGGAGTATGCTTGAAGAGATACCTGTTCCATGCCGACACAGAAAGTTAAGC 233

DB 3108 AGTTAAGTACGATAGCCTAGGAGATACACTCTTACCATGCCGAAACAGAAAGTTAAGC 3049
QY 234 TTCAGACCGCCGATAGTATGGGGATCGCC 267
DB 3048 CCTAGAACCGCGAAGTAGTTGGGGGTTGCC 3015

RESULT 3

US-10-158-844-53/c
; Sequence 53, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
; APPLICANT: Kunsch et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/158.844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB340P1D1
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3510 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-158-844-53

Query Match 43.3%; Score 115.6; DB 16; Length 3510;
Best Local Similarity 66.8%; Pred. No. 1.1e-26;
Matches 183; Conservative 0; Mismatches 84; Indels 7; Gaps 1;

QY 1 TATATGGAAGTAAAGCCCTTGAGAGATGATCAGGTAGCTAGGAGTGAAGTGAAGCCCG 60
DB 3288 TATATATCAGTAGAGCCCTTGAGAGATGATCAGGTAGCTAGGAGTGAAGTGAAGTGG 3229
QY 61 TGAGCGGTGAGAGCGGACGATCTACTATCGTGAAGACTTAACCAAGTCAACAGTAGT 120
DB 3228 CGACACATGTAGCGGACGATCTACTATCGTGAAGACTTAACCAAGTCAACAGTAGT 3169
QY 121 TGTTCGAGATATTAATTAATATCTAGTTTGAAGGAGAG-----TTCCTTAT 173
DB 3168 TGAAGCGGACGCTTTCTTAATTAATGAATAGATTAATTTTGAAGTATTAAGTACG 3109
QY 174 AGTGTGGGAGTATGCTTGAAGAGATACCTGTTCCATGCCGACACAGAAAGTTAAGC 233
DB 3108 AGTTAAGTACGATAGCCTAGGAGATACCTGTTCCATGCCGACACAGAAAGTTAAGC 3049
QY 234 TTCAGACCGCGATAGTATGGGGATCGCC 267
DB 3048 CCTAGAACCGCGAAGTAGTTGGGGGTTGCC 3015

QY 61 TGAGCGGTGAGCGGACCACTACTATTCGTCGAGACTTAAACCAAGTCAACAGTACT 120
DB 3228 CGACACATGATGCGGACCTAATATAGCTCGAGACTTATCCAAAGTACTGAGATA 3169
QY 121 TGTTCGAGATTAATTGAATATATCTAGTTTGAAGGAGAGAG-----TTCTCTTAT 173
DB 3168 TGAAGCGAAGCGTTTCTTTTAAATTGAATAGATTCAATTTAGTAGGATTAAGTACAG 3109
QY 174 AGTGTGTGGCGATACCTTGAAGATACACTGTTCCATGCGCAACACAGAAATTAGC 233
DB 3108 AGTTAAGTGAAGATAGCCTAGAGATACACTGTACCCATGCGCAACACAGAAATTAGC 3049
QY 234 TTCAGCAGCGCGATAGTAGTGGGAGATCGCCCC 267
DB 3048 CCTAGACCGCGGAAGTAGTTGGGGTTGCCCCC 3015

RESULT 2
US-08-961-527-16
Sequence 16, Application US/08961527
Patent No. 6420135

GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 8411 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-16

Query Match 43.3%; Score 115.6; DB 4; Length 8411;
Best Local Similarity 66.8%; Pred. No. 1.3e-31;
Matches 183; Conservative 0; Mismatches 84; Indels 7; Gaps 1;

QY 1 TATATGAGATTAAGACCCCTGAGAGATGATCAGGTAGTGGCTGGAAGTACGAGCCGC 60
DB 316 TATATATCAGTAAGAGCCCTGAGAGATGATCAGGTAGTGAAGTGAAGTGTGG 375
QY 61 TGAGCGGTGAGCGGACCACTACTATCGGTGAGAGCTTAACCAAGTCAACAGTACT 120
DB 376 CGACACATGATGCGGACCTAATCTAATAGCTGAGAGCTTATCCAAAGTAACTGAGATA 435
QY 121 TGTTCGAGATTAATTGAATATATCTAGTTTGAAGGAGAG-----TTCTCTTAT 173

DB 436 TGAAGCGAAGCGTTTCTTAATTAATGATATATTCATTTTGAAGTAGTATTACTGAG 495
QY 174 AGTGTGTGGCGATAGCTTGAAGATACACTGTTCCATGCGCAACAGAAATTAGC 233
DB 496 AGTTAAGTGAAGATAGCTTGAAGATACACTGTTCCATGCGCAACAGAAATTAGC 555
QY 234 TTCAGCAGCGCGATAGTAGTGGGAGATCGCCCC 267
DB 556 CCTAGACCGCGGAAGTAGTTGGGGTTGCCCCC 589

RESULT 3
US-08-961-527-96
Sequence 96, Application US/08961527
Patent No. 6420135

GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 11915 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-96

Query Match 43.3%; Score 115.6; DB 4; Length 11915;
Best Local Similarity 66.8%; Pred. No. 1.6e-31;
Matches 183; Conservative 0; Mismatches 84; Indels 7; Gaps 1;

QY 1 TATATGAGATTAAGACCCCTGAGAGATGATCAGGTAGTGGCTGGAAGTACGAGCCGC 60
DB 255 TATATATCAGTAAGAGCCCTGAGAGATGATCAGGTAGTGAAGTGAAGTGTGG 314
QY 61 TGAGCGGTGAGCGGACCACTACTATCGGTGAGAGCTTAACCAAGTCAACAGTACT 120
DB 315 CGACACATGATGCGGACCTAATCTAATAGCTGAGAGCTTATCCAAAGTAACTGAGATA 374
QY 121 TGTTCGAGATTAATTGAATATATCTAGTTTGAAGGAGAG-----TTCTCTTAT 173
DB 375 TGAAGCGAAGCGTTTCTTAATTAATGAATATTAATCAATTTTGAAGTAACTGAG 434
QY 174 AGTGTGTGGCGATAGCTTGAAGATACACTGTTCCATGCGCAACAGAAATTAGC 233
DB 435 AGTTAAGTGAAGATAGCTTGAAGATACACTGTTCCATGCGCAACAGAAATTAGC 494

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 / Search time 517.964 Seconds
(without alignments)
2705.971 Million cell updates/sec

Title: US-10-088-666-1

Perfect score: 267
1 tatatggaagtaagaccctc.....agtagtcgggggacgccccc 267

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002s:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003s:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	100.0	267	5	AAF61566 Lactobac
2	156.4	88.6	317	5	AAF61572 L. coryni
3	156.4	88.6	317	5	AAF61571 L. coryni
4	143.6	53.8	326	5	AAF61567 Lactobac
5	139.4	52.2	335	5	AAF61574 Pediococ
6	138.2	51.8	326	5	AAF61575 Pediococ
7	136.4	51.1	336	5	AAF61573 Lactobac
8	131	49.1	110000	10	ADF77343_05
9	131	49.1	110000	10	ADF77343_07
10	131	49.1	110000	10	ADF77343_18
11	121.6	45.5	110000	9	ADB12064_00
12	120	44.9	110000	9	ADB12064_05
13	120	44.9	110000	9	ADB12064_14
14	119.2	44.6	338	5	AAF61570 Lactobac
15	118.8	43.7	110000	9	AAF61570 Lactobac
16	118.8	43.4	31702	6	ABQ67190 Listeria
17	118.6	43.3	3510	2	AAV52186 Streptoc
18	115.6	43.3	3989	2	AAV52186 Streptoc
19	115.6	43.3	3989	6	ABQ67194 Listeria
20	115.6	43.3	5048	3	AAAG5735 Enterococ
21	115.6	43.3	5048	6	ABK15101 DNA encod

22	115.6	43.3	8411	2	AAV52149	AAV52149 Streptoc
23	115.6	43.3	11915	2	AAV52229	AAV52229 Streptoc
24	115.6	43.3	110000	10	ABQ56454_16	Continuation (17 o
25	115.6	43.3	110000	10	ABQ56454_17	Continuation (18 o
26	115.6	43.3	110000	10	ABQ56454_18	Continuation (19 o
27	115.6	43.3	110000	10	ABQ56454_19	Continuation (20 o
28	114.2	42.8	1389	6	ABQ70505	Continuation (21 o
29	114.2	42.8	1389	6	ABQ70505	Continuation (22 o
30	114.2	42.8	5349	6	ABQ70982	Abq70982 Listeria
31	114.2	42.8	110000	6	ABQ67196_0	Abq67196 Listeria
32	114.2	42.8	110000	6	ABQ67196_5	Continuation (6 of
33	114.2	42.8	110000	6	ABQ69245_02	Continuation (3 of
34	114.2	42.8	110000	6	ABQ69245_17	Continuation (18 o
35	114.2	42.8	110000	6	ABQ69245_18	Continuation (19 o
36	114.2	42.8	110000	6	ABQ69245_19	Continuation (20 o
37	114.2	42.8	110000	6	ABQ69245_24	Continuation (25 o
38	114.2	42.8	110000	6	ABQ69245_27	Continuation (28 o
39	114.2	42.8	110000	6	ABQ67197_01	Continuation (2 of
40	114.2	42.8	110000	6	ABA03041_02	Continuation (3 of
41	114.2	42.8	110000	6	ABA03041_24	Continuation (25 o
42	114.2	42.8	110000	6	ABA03041_26	Continuation (27 o
43	114.2	42.8	319630	6	ABQ67194	Abq67194 Listeria
44	114	42.7	9797	2	AAV52149	AAV52149 Streptoc
45	114	42.7	9797	6	ABQ59282	ABQ59282 Enterococ

ALIGNMENTS

RESULT 1
ID AAF61566 standard; DNA; 267 BP.
AC AAF61566;
DT 02-JUL-2001 (first entry)
XX Lactobacillus brevis 23S rRNA-spacer-5S rRNA DNA fragment.
XX Lactobacillus brevis 23S rRNA; detection; probe; brewing; beer; contamination; ss.
XX Lactobacillus brevis.
OS Lactobacillus brevis.
XX DE19945964-A1.
XX 05-APR-2001.
XX 24-SEP-1999; 99DE-01045964.
XX 24-SEP-1999; 99DE-01045964.
XX PR 24-SEP-1999; 99DE-01045964.
XX PA (BIOT-) BIOTECON DIAGNOSTICS GMBH.
XX Pandke M, Gasch A, Berghof K;
XX WPI; 2001-246136/26.
XX PT Detecting contaminating microorganisms in brewing, by nucleic acid
PT amplification and hybridization, either non-specific or genus- or species
PT -specific.
XX Claim 9(1); Page 9; 48pp; German.
XX PS
XX This invention describes a novel method for detecting microorganisms (A)
XX of importance in brewing which comprises treating a sample with at least
XX two primers (P1) that hybridize to a consensus region in the nucleic acid
XX of (A), at least part of the microbial nucleic acid is amplified, the
XX amplicon is treated with a sequence common to all (A) or specific for one or
XX more families, genera or species, and any formation of hybrids is
XX detected. The method is used to detect, identify and/or characterize
XX microorganisms in beer or brewing materials, particularly for detecting
XX contamination. The method may detect the entire range of contaminating

CC microbes, either as a general test for contamination or as a test
CC specific for particular genera or (sub)species. It is quicker than known
CC microbiological methods, and can detect several organisms in the same
CC sample, including organisms not presently recognised as contaminants. The
CC method provides an early indication of contamination and can be automated
CC for high throughput analysis

SQ Sequence 267 BP; 77 A; 51 C; 77 G; 62 T; 0 U; 0 Other;

Query Match	Score	DB	Length
100.0%	267	5	267

```
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

[illegible]

RESULT 2	
AAF61572	
ID	AAF61572 standard; DNA; 317 BP.

AC AAF61572;

DT 02-JUL-2001 (first entry)

DE L. coryniformis ssp torquens 23S rRNA-spacer-5S rRNA DNA fragment

KW 23S rRNA; 5S rRNA; detection; probe; brewing; beer; contamination; SS

Lactobacillus coryniformis.

PN DE19945964-A1

PD 05-APR-2001.

PR 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PA (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K;

DR WPI; 2001-246136/26.

PT Detecting contaminating microorganisms in brewing, by nucleic acid amplification and hybridization, either non-specific or genus- or species-specific.

PS Claim 9(1); Page 10; 48bp; German

AA This invention describes a novel method for detecting microorganisms (A)
CC of importance in brewing which comprises treating a sample with at least
CC two primers (P1) that hybridize to a consensus region in the nucleic acid
CC of (A), at least part of the microbial nucleic acid is amplified, the
CC amplicon is treated with at least one probe (P2) that hybridizes

CC specifically with a sequence common to all (a) or specific for one or
CC more families, genera or species, and any formation of hybrids is
CC detected. The method is used to detect, identify and/or characterize
CC microorganisms in beer or brewing materials, particularly for detecting
CC contamination. The method may detect the entire range of contaminating
CC microorganisms, either as a general test for contamination or as a test
CC specific for particular genera or (sub) species. It is quicker than known
CC microbiological methods, and can detect several organisms in the same
CC sample, including organisms not presently recognized as contaminants. The
CC method provides an early indication of contamination and can be automated
CC for high throughput analysis

50 Sequence 317 BP; 90 A; 59 C; 96 G; 72 T; 0 U; 0 Other;

Query Match 58.6%; Score 156.4; DB 5; Length 317;

Matches 216; Conservative 0; Mismatches 51; Indels 9; Gaps 2

QY 1 TAAATGGAAGTAAAGCCCTGAGAGATATACAGGTAGATAGGCTGAAATAGACGCCG 60
Db 24 TTTATGAAAGTAAAGCCCTGAGAGATATACAGGTAGATAGGTTGGAAGTGAAGTCCG 83
QY 61 TGAAGCGTGAAGCGGACCAAGTCTAATCGTGTGAGGACTTAACCAAG----TCACAC 115
Db 84 TGAAGCATGAAGCGGACCAATCTATATCGTGTGAGGACTTAACCAAGTAGACATGATG 143
QY 116 GTAGTGTGTTGAGAAATATTGAATTAATATCTAGTTTGAAGG---AAGAATCTCTT 171
Db 144 GTGTTAGTTTAAAGGCGCAAGAAATGAATATCCAGTTTGAAGGCGCAAGTTCTCAGAAA 203
QY 172 ATAGTGTGAGCGATAGGCTGGAAGGATACACTGTGTTCCATGCCGCAACACAGAAAGTTAA 231
Db 204 GTGTGTGAGGAGATAGCAAGAAAGATACACTGTGTTCCATGTGCAACACAGAAAGTTAA 263
QY 232 GCTTCAGCACGCCGAGATAGTATTGGGGGATGCCCC 267
Db 264 GCTTCCTTAAGCGCGAAGTAGTGTGGGGAGCACCCC 299

RESULT 3
AAF61571
ID AAF61571 standard; DNA; 317 BP

AC AAF61571

DT 02-JUL-2001 (first entry)

DE L. coryniformis bsp coryniformis 23S rRNA-spacer-5S rRNA DNA fragment.

23S rRNA; detection; probe; brewing; beer; contamination; ss-

Lactobacillus coryniformis OS

PN DE19945964-A1.

PD 05-APR-2001

PF 24-SEP-1999; 99DE-01045964.

PR 24-SEP-1999; 99DE-01045964.

PA (BIOT-) BIOTECON DIAGNOSTICS GMBH.

PI Fandke M, Gasch A, Berghof K,

DR WPI; 2001-246136/26.

PT Detecting contaminating microorganisms in brewing, by nucleic acid
PT amplification and hybridization, either non-specific or genus- or species
PT -specific.

PS Claim 9(i); Page 10; 48pp; German

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 11:05:13 ; Search time 2085.49 Seconds

(without alignments)
6054.391 Million cell updates/sec

Title: US-10-088-666-1

Perfect score: 267 1 tatatcgagaagtaagaccccc.....agtagtcggggagtcacccc 267

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1:
1: gb_ba:*
2: gb_hig:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pal:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Fred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	100.0	267	6	AX105729 Sequence
2	173.4	64.9	269050	1	AL935255 Lactobaci
3	173.4	64.9	297050	1	AL935258 Lactobaci
4	171.8	64.3	298050	1	AL935260 Lactobaci
5	171.8	64.3	298050	1	AL935261 Lactobaci
6	171.8	64.3	324050	1	AL935253 Lactobaci
7	156.4	58.6	317	6	AX105734 Sequence
8	156.4	58.6	317	6	AX105735 Sequence
9	143.6	53.8	326	6	AX105730 Sequence
10	139.4	53.2	335	6	AX105737 Sequence
11	138.2	51.8	326	6	AX105738 Sequence
12	136.4	51.1	336	6	AX105736 Sequence
13	131	49.1	177811	1	AE017206 Lactobaci
14	131	49.1	183043	6	AX326717 Sequence
15	131	49.1	300478	1	AE017201 Lactobaci
16	131	49.1	300886	1	AE017205 Lactobaci
17	131	49.1	349980	6	AX326712 Sequence
18	131	49.1	349980	6	AX326713 Sequence
19	119.2	44.6	338	6	AX105733 Sequence

C	20	115.8	43.4	31702	6	AX413012 Sequence
C	21	115.6	43.3	3510	6	C0788962 Sequence
C	22	115.6	43.3	3510	6	AR218821 Sequence
C	23	115.6	43.3	3510	6	BD003733 Polynucle
C	24	115.6	43.3	3989	6	BD193765 Enterococ
C	25	115.6	43.3	5048	6	BD263589 Novel str
C	26	115.6	43.3	5048	6	AX343071 Sequence
C	27	115.6	43.3	8411	6	C0788925 Sequence
C	28	115.6	43.3	8411	6	AR218784 Sequence
C	29	115.6	43.3	8411	6	BD003696 Polynucle
C	30	115.6	43.3	9873	1	AE007481 Streptoco
C	31	115.6	43.3	10058	1	AE007489 Streptoco
C	32	115.6	43.3	10551	1	AE007319 Streptoco
C	33	115.6	43.3	10557	1	AE008537 Streptoco
C	34	115.6	43.3	10856	1	AE008546 Streptoco
C	35	115.6	43.3	11385	1	AE007495 Streptoco
C	36	115.6	43.3	11660	1	AE008385 Streptoco
C	37	115.6	43.3	11915	6	C0789005 Sequence
C	38	115.6	43.3	11915	6	AR218864 Sequence
C	39	115.6	43.3	11915	6	BD003776 Polynucle
C	40	115.6	43.3	137952	2	SPNEU1909 Streptoco
C	41	115.6	43.3	232807	2	SPNEU1901 Streptoco
C	42	115.6	43.3	301488	1	AE016950 Enterococ
C	43	115.6	43.3	304454	1	AE016955 Enterococ
C	44	115.6	43.3	349980	6	AX571765 Sequence
C	45	115.6	43.3	349980	6	AX571766 Sequence

ALIGNMENTS

RESULT 1
AX105729
LOCUS AX105729 Sequence 1 from Patent WO0123605.
DEFINITION AX105729
ACCESSION AX105729
VERSION AX105729.1 GI:13921742
KEYWORDS
SOURCE Lactobacillus brevis
ORGANISM Lactobacillus brevis
Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
Lactobacillus.

REFERENCE

1 Pandke, M., Gasch, A. and Berghof, K.
Method and nucleic acids for determining the presence of
micro-organisms specific to the brewing process
Patent: WO 0123605-A 1 05-APR-2001;
Biotecon Diagnostics GmbH (DE)
Location/Qualifiers

FEATURES

1..267
/organism="Lactobacillus brevis"
/mol_type="unassigned DNA"
/db_xref="taxon:1580"

ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 6-6e-72;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TATATGGAAGTAAGACCCCTGAGAGATGATCAGGTGATAGGCTGGAAGTAGAGCGCCG 60
1 TATATGGAAGTAAGACCCCTGAGAGATGATCAGGTGATAGGCTGGAAGTAGAGCGCCG 60
Db 1 TATATGGAAGTAAGACCCCTGAGAGATGATCAGGTGATAGGCTGGAAGTAGAGCGCCG 60
QY 61 TGAAGCGTGGAGGAGGACCACTAATTAATCGGTGAGAGCTTAACCAAGTCAACAGTAGT 120
61 TGAAGCGTGGAGGAGGACCACTAATTAATCGGTGAGAGCTTAACCAAGTCAACAGTAGT 120
QY 121 TGTTCGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
121 TGTTCGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
Db 121 TGTTCGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
QY 181 TGGCGATAGCGCTGAAGATACACCTTCCCAAGCCGGAACAGAGTTAAGCTTAGACA 240
181 TGGCGATAGCGCTGAAGATACACCTTCCCAAGCCGGAACAGAGTTAAGCTTAGACA 240

Ds 161 TGGCGATAGCCTGAAGATACCTGTTCCTCCATGCCGACACAGATTAAGCTTCAGCA 240
Oy 241 CGCCGATAGTATGGGGATGCCCC 267
Ds 241 CGCCGATAGTATGGGGATGCCCC 267

RESULT 2
AL935255 269050 bp DNA linear BCT 21-FEB-2003
LOCUS Lactobacillus plantarum strain WCFS1 complete genome; segment 4/11.
DEFINITION AL935255 AL935253
ACCESSION AL935255.1 GI:28270709
VERSION
KEYWORDS
SOURCE
ORGANISM
Lactobacillus plantarum WCFS1
Lactobacillus plantarum WCFS1
Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
Lactobacillus.

REFERENCE
AUTHORS
1 Kieerebezen, M., Boekhorst, J., van Kranenburg, R., Molenaar, D.,
Kulbers, O.P., Leer, R., Tarchini, R., Peters, S.A., Sandbrink, H.M.,
Peters, M.W.E.U., Stiekema, W., Lankhorst, R.V.K., Bron, P.A.,
Hoffer, S.M., Groot, M.N.N., Kerkhoven, R., de Vries, M., Ursing, B., de
Vos, W.M. and Siezen, R.J.
Complete genome sequence of Lactobacillus plantarum WCFS1
Proc. Natl. Acad. Sci. U.S.A. 100 (4), 1990-1995 (2003)
2 (bases 1 to 269050)
Kieerebezen, M. and Siezen, R.J.
Direct Submission
Submitted (14-OCT-2002) Wageningen Centre for Food Sciences, P.O.
Box 557, 6700 AN Wageningen, The Netherlands
FEATURES
source
1. 269050
/organism="Lactobacillus plantarum WCFS1"
/mol_type="genomic DNA"
/strain="WCFS1"
/db_xref="taxon:220668"
/note="single colony isolate of strain NCIM8826"
596. 2608
/gene="ekt2"
/locus_tag="lp_1083"
596. 2608
/gene="ekt2"
/locus_tag="lp_1083"
/EC_number="2.2.1.1"
/codon_start=1
/transl_table=11
/product="transketolase"
/protein_id="CAD63615.1"
/db_xref="GI:28270710"
/db_xref="GOA:Q88XU3"
/db_xref="TrEMBL:Q88XU3"
/translation="MQAPDEIDELAVNRIYLSAEMIRAGSGHGLPLGAAPMMML
WSRQLRVDPQSQMANRDRFVLSAGHSHLYSLHVGSRNTMGELOPFQFSGHNP
GHEPGLVDGVAATGTPGQGGAGVAGWALAEKQLAOFSSPOLVDHITVALVGGD
LMEGVSHBAFAGNQQOLKLVLDNVAISLDGTPSFRDVAORPASGWDITQV
GGNDITAINAIBAKHSSRSPLIAVTTTGGAPQAGTAVHGNPAGTGLTSRQ
OLGMOAPPTVAAPAKKQOLIORGQRRWQSGREALPOAOFHFEQOLSMRIP
VPLATSLPFDGGAQASRTSHTVOKLAOLPGLVGAADLASNKTTIRDOOLSTP
VTPDORNAFGRFRGKGTILNGMLHGLRFRGGTPLVPSDYRAAIRLAAMKLV
TVFTHDSTAVGEDGTHPIOLSLRIPLVITIRADNRETYAAMQALNATDH
TVLVITRQNLAVLSGSKTAMTGVARGVLSPOGRPMGILASGSEVALAQQA
KATLGEDVAVSMPSPMENFASSTNYQAVLPQIRRVAVEMGATSGWERYVGLDG
AVGLDHFASGAADQVLAANOFTVADLVRTYQOTVHDLGQQLRVI"

gene
CDS

2628. 3521
/gene="arod1"
/locus_tag="lp_1084"
2628. 3521
/gene="arod1"
/locus_tag="lp_1084"
/EC_number="1.1.1.25"

gene
CDS

/codon_start=1
/transl_table=11
/product="snkimate 5'-dehydrogenase"
/protein_id="CAD63616.1"
/db_xref="GI:28270711"
/db_xref="GOA:Q88XU2"
/db_xref="TrEMBL:Q88XU2"
/translation="MIDGQALFPIAHARHSRSPRNHNSFDYGMANRYALADVA
PGLAGTLAGIRAGLHAGLNLNMPKQIVPLDGLTYAORIQAVNTIKBAGLIG
DSTDAGVQDILHOGILNYPKHVWAGAGAGKALIAALDYAVATGPKRNNRY
ADVQMLBOLSTSTVPLBLCAVEQVRYKQYARISNLIINATNGMTTACGAPLISI
IQOLVPOQTVYDVIIYQPLETFLRLARQGGCTTNGLGLMLIQGLALAEPTGKMPV
QGVRODMLASIQATPPIITAQA"
3562. 4557
/gene="arod1"
/locus_tag="lp_1085"
3562. 4557
/gene="arod1"
/locus_tag="lp_1085"
/EC_number="4.1.2.15"
/EC_number="5.4.99.5"
/codon_start=1
/transl_table=11
/product="phospho-2-dehydro-3-deoxyheptonate aldolase /
choistmate mutase"
/protein_id="CAD63617.1"
/db_xref="GI:28270712"
/db_xref="GOA:Q88XU1"
/db_xref="TrEMBL:Q88XU1"
/translation="MIIILKPAATATAMQAKLVSOFGADHEIFTHNRIALQGLREAL
PMDQASSTITRANPAAYOSRLFPBETITIKTPHSAVIGNCFVMMAGPQSVSEAE
VPMQAAVAVQAGATTVRGGAFFKRTSPYSFQGLGETGQOYRTADNKGIDMTETW
DERVANVATYDIPOIGARNMNFSLKAVGQTOIPLAKRMSATIDIVNAAYIA
AGNHQIMLEBERGIRTFENKRTNTEFDLGAVALQSLTHYPTVPSHAGVETWELTP
MAMAGVAAAGSAGIVEIHDPQALSDGQALKPATYIKMAQOAFQLHDMQTM"
4561. 5646
/gene="arob"
/locus_tag="lp_1086"
4561. 5646
/gene="arob"
/locus_tag="lp_1086"
/EC_number="4.6.1.3"
/codon_start=1
/transl_table=11
/product="3-dehydroquinate synthase"
/protein_id="CAD63618.1"
/db_xref="GI:28270713"
/db_xref="GOA:Q88XU0"
/db_xref="TrEMBL:Q88XU0"
/translation="MPMTNIEVKTIRYKQYQIKIMINSILHLGLTVTKYCAQVAVT
DNYGQYVAKVYDDELTAAGFVRVWTPVAGSEKSMQVGSLLIOLSAHRSRSDGV
LALGGVVDLAFVASTYMRGIALIQVPTSLAOVDSVGGKTAIDLPSTGNLVSF
YQDLVAVDPAIIVLPRLABGEGEIVKCALVGGDPQWSLHQTSTVAALTPAAPD
RLAASVAFARVVMADHESQORQLNFGHTIGHAVELADQQLHGEHVAIGLVQC
LIPAAHGLAPTSITSLKARLNAVGLPTLPIDAVAAVQHDKXGHTALTWVYL
SAYGPHLPIAVATDLATWMDLMSVS"
complement (6078. 7499)
/locus_tag="lp_1087"
/locus_tag="lp_1087"
/codon_start=1
/transl_table=11
/product="cation transport protein"
/protein_id="CAD63619.1"
/db_xref="GI:28270714"
/db_xref="TrEMBL:Q88XU9"
/translation="MTRLOSQVNTQOITWPLIIGLITWLSAPPRPAGLSLHMHMLAF
VGTIGITQPLPIAGVAILGLVGLVMTWMAAGFNTWMIIMIAFLSRGF
VKTGLRRVALVFLVFKKKTGLGVALVGLVSPSPSTARAGIVFPIISLA
KAYGSPQNGTQKSGSLIMSFGHNIATSGLFMTAPNVAIVAAQAMQIHITMM
DWLLAALIPGLICLISVPAIYWLVPPEIKDIPNAAVANELEQGAATLAKKIMH
ITMLALPMSASLTGLDPAIVAFIVATLLITGTSVDLIDHENGANTLWFSILI
FMAGKLSIGFTPLMSQITWGLQGISMLVLAALMLVYFYSHYLPAGNTAVSANYG

gene
CDS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2004, 12:43:29 ; Search time 2084.68 Seconds
(without alignments)
6056.719 Million cell updates/sec

Title: US-10-088-666-1

Sequence: 1 tatatggaagaaagacccct.....agtagtcggggatcgcgcc 267

Scoring table: OLIGO-NUC
Gapop 60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word-size: 10

Total number of hits satisfying chosen parameters: 1357868

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 100 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_hcg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pac:*

7: gb_pl:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_ets:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267	100.0	267	6 AX105129	Sequence
2	92	34.5	117	1 LBRNRS	X02026 Lactobacilli
3	73	27.3	326	6 AX105730	Sequence
4	50	18.7	338	6 AX105733	Sequence
5	50	18.7	351	6 AX105731	Sequence
6	50	18.7	414	6 AX105732	Sequence
7	50	18.7	1541	1 AR098107	Lactobacilli
8	44	16.5	249	6 AR397133	Sequence
9	44	16.5	3989	6 BD193765	Enterococ
10	44	16.5	9797	6 BD193956	Enterococ
11	44	16.5	22960	6 BD193751	Enterococ
12	44	16.5	176545	1 AE016957	Enterococ
13	44	16.5	177911	1 AE017206	Lactobacilli
14	44	16.5	183043	6 AX926717	Sequence
15	44	16.5	300478	1 AE017201	Lactobacilli
16	44	16.5	300886	1 AE017205	Lactobacilli
17	44	16.5	301488	1 AE016950	Enterococ
18	44	16.5	304454	1 AE016956	Enterococ
19	44	16.5	326434	1 AE016947	Enterococ

20	44	16.5	349980	6 AX926712	Sequence
21	44	16.5	349980	6 AX926713	Sequence
22	40	15.0	317	6 AX105734	Sequence
23	40	15.0	317	6 AX105735	Sequence
24	40	15.0	326	6 AX105738	Sequence
25	40	15.0	335	6 AX105737	Sequence
26	40	15.0	2942	6 BD107241	Sequence
27	37	13.9	357	1 STRP20	Sequence
28	37	13.9	576	1 S43413	Sequence
29	36	13.5	300050	1 AP004595	Sequence
30	35	13.1	54	6 AR358918	Sequence
31	35	13.1	107	1 MCRN01	Sequence
32	35	13.1	107	1 MCRN01	Sequence
33	35	13.1	109	1 MYCRGAF3	Sequence
34	35	13.1	115	1 STARRASA	Sequence
35	35	13.1	115	1 STARRASE	Sequence
36	35	13.1	187	6 AR358564	Sequence
37	35	13.1	216	6 AR358471	Sequence
38	35	13.1	221	6 AR358454	Sequence
39	35	13.1	235	6 AR358418	Sequence
40	35	13.1	242	6 AR358420	Sequence
41	35	13.1	283	6 AR358342	Sequence
42	35	13.1	295	6 AR358322	Sequence
43	35	13.1	300	6 AR358297	Sequence
44	35	13.1	309	6 AR358208	Sequence
45	35	13.1	327	6 AR358224	Sequence
46	35	13.1	330	6 AR358206	Sequence
47	35	13.1	338	6 AR358128	Sequence
48	35	13.1	340	6 AR358077	Sequence
49	35	13.1	348	6 AR358057	Sequence
50	35	13.1	361	6 AR357829	Sequence
51	35	13.1	361	6 AR358178	Sequence
52	35	13.1	367	6 AR357941	Sequence
53	35	13.1	367	6 AR358101	Sequence
54	35	13.1	370	6 AR358152	Sequence
55	35	13.1	371	6 AR358000	Sequence
56	35	13.1	386	6 AR357946	Sequence
57	35	13.1	389	6 AR357772	Sequence
58	35	13.1	400	6 AR357493	Sequence
59	35	13.1	400	6 AR357506	Sequence
60	35	13.1	400	6 AR357513	Sequence
61	35	13.1	400	6 AR357516	Sequence
62	35	13.1	400	6 AR357520	Sequence
63	35	13.1	400	6 AR357527	Sequence
64	35	13.1	400	6 AR357542	Sequence
65	35	13.1	400	6 AR357543	Sequence
66	35	13.1	400	6 AR357552	Sequence
67	35	13.1	400	6 AR357557	Sequence
68	35	13.1	400	6 AR357601	Sequence
69	35	13.1	400	6 AR357620	Sequence
70	35	13.1	400	6 AR357624	Sequence
71	35	13.1	400	6 AR357626	Sequence
72	35	13.1	400	6 AR357630	Sequence
73	35	13.1	400	6 AR357650	Sequence
74	35	13.1	400	6 AR357740	Sequence
75	35	13.1	400	6 AR357748	Sequence
76	35	13.1	400	6 AR357816	Sequence
77	35	13.1	400	6 AR357911	Sequence
78	35	13.1	401	6 AR357613	Sequence
79	35	13.1	425	6 AR357596	Sequence
80	35	13.1	458	6 AR357639	Sequence
81	35	13.1	475	6 AR357503	Sequence
82	35	13.1	587	6 AR357436	Sequence
83	35	13.1	596	6 AR357478	Sequence
84	35	13.1	840	6 AR354390	Sequence
85	35	13.1	1051	6 AR357456	Sequence
86	35	13.1	1089	1 AF270293	Sequence
87	35	13.1	1089	6 AR486247	Sequence
88	35	13.1	1089	6 AX145611	Sequence
89	35	13.1	1171	6 AR357421	Sequence
90	35	13.1	1404	1 MPCRNRN3	Sequence
91	35	13.1	2209	6 AR357434	Sequence
92	35	13.1	2662	1 AF270294	Staphyloc

